DOCUMENT RESUME

ED 153 156 CG 012 508

AUTHOR	Guess, L. Lynn; Tuchfeld, Barry S.
TITLE	Manual for Drug Abuse Treatment Program Self-Evaluation. Treatment Program Monograph Series No. 1.
INSTITUTION	Research Triangle Inst., Durham, N.C.
SPONS AGENCY	National Inst. on Drug Abuse (DHEW/PHS), Rockville, Md.
PUB DATE	77
NOTE	70p.; For related document see CG 012 120 ; Parts of document may be marginally legible due to light type
AVAILABLE FROM	Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 (Stock number 017-024-00572-0, price $1.60)
EDRS PRICE	MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS	Administrative Personnel; *Clinics; Data Analysis; *Drug Abuse; *Organizational Effectiveness; *Program Evaluation; Program Guides; *Self Evaluation

ABSTRACT

This monograph is the first of a series produced by the National Institute on Drug Abuse. The series will provide simple, straightforward presentations relevant to the operational aspects of the clinical setting. This monograph provides the basic tools for each program to establish reasonable treatment goals and assess the quality of its service by measuring the progress of the program in reaching those goals. Such evaluation data can be used by the program manager to upgrade treatment services and demonstrate accountability for funding. This manual describes the nature and purpose of clinical self-evaluation, the limitations on the use of the data, and methods for conducting such evaluations using data collection procedures already in operation in the program. It should be a useful resource for treatment personnel. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED153156

NATIONAL INSTITUTE ON DRUG ABUSE

TREATMENT PROGRAM MONOGRAPH SERIES

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

MANUAL FOR DRUG ABUSE TREATMENT PROGRAM SELF-EVALUATION

CG 012508

U.S. DEPARTMENT OF HEALTH EDUCATION AND WELFARE • Public Health Service • Alcohol, Drug Abuse, and Mental Health Administration

MANUAL FOR DRUG ABUSE TREATMENT PROGRAM SELF-EVALUATION

By

L. Lynn Guess
Barry S. Tuchfeld

Research Triangle Institute

National Institute on Drug Abuse
11400 Rockville Pike
Rockville, Maryland 20852

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402 - Price $1.60
Stock No. 017-024-00572-0

This volume, part of a Treatment Program Monograph Series, was prepared for the National Institute on Drug Abuse by Research Triangle Institute, Post Office Box 12194, Research Triangle Park, North Carolina 27709, under Contract Number 271-75-1016, Work Order Number 3.

FOREWORD

The National Institute on Drug Abuse is developing a series of publications designed for the treatment program manager in the field. Dealing with the day-to-day problems encountered by program managers in the areas of evaluation, followup studies, and treatment, etc., the Treatment Program Monograph Series will attempt to provide simple, straightforward presentations relevant to the operational aspects of the clinical setting.

This first monograph in the series, a Manual for Drug Abuse Treatment Program Self-evaluation, provides the basic tools for each program to establish reasonable treatment goals and assess the quality of its service by measuring the progress of the program in reaching those goals. Such evaluation data can be used by the program manager to upgrade treatment services and demonstrate accountability for funding.

This manual describes the nature and purpose of clinical self-evaluation, the limitations on the use of the data, and methods for conducting such evaluations using data collection procedures already in operation in the program.

We hope that this monograph series becomes an important resource for treatment personnel and that it will speak to the needs of the treatment community.

Robert L. DuPont, M.D.
Director
National Institute on Drug Abuse

CONTENTS

		Page
	FOREWORD	iii
Part 1	INTRODUCTION AND EXPLANATION OF SELF-EVALUATION	1
I.	Introduction and Overview	1
	Introduction	1
	Why Bother?	1
	Questions and Answers	2
	Clinical Staff Support	3
	Useful Results	3
II.	The Process of Evaluation	6
	Defining Goals and Objectives	6
	Summary of the Evaluation Process	8
III.	Important Points to Consider	9
	Keeping It Simple	9
	Building and Using a Data Base	9
	The Real vs. the Ideal	10
Part 2	SUGGESTED PROGRAM MEASURES AND COMPUTATIONAL INSTRUCTIONS	11
IV.	Introduction to the Measures	11
	Program Measures	11
	The Element of Time	14
V.	Computation of the Measures	15
	Introduction	15
	Employment Measures	17
	Measures of Drug Use	21
	Criminality Measures	27
	Termination Measures	28
	Other Measures	33
Part 3	COMPARISON GROUPS AND PROCEDURES FOR SELF-EVALUATION	37
VI.	Subgroups and Cohorts	37
	Subgroup Distinctions	37
	Constructing and Comparing Subgroups	38
	Cohorts: Time-Based Groups	43
	Computing Measures for Cohorts	47
VII.	Trends	51
	Introduction	51
	Analyzing Trends	52
	Trends for Cohort Values	53
VIII.	Additional Procedures	55
	Dealing With Small Groups	55
	Some Further Suggestions	56

Page

SUGGESTED READINGS 59

APPENDIX A SAMPLE WORKSHEET FOR DARP FORMULAS. A-1

APPENDIX B SAMPLE WORKSHEET FOR CODAP FORMULAS B-2

Figures

1 Determining normal treatment outcomes and identifying problems. 4

2 Comparing modality-environments to help determine resource allocation. 4

3 Comparing progress of groups of clients admitted to treatment during the same months (Cohorts). 5

4 Relationship between goals, objectives, and criterion measures. 6

5 Summary of evaluation process 7

6 Sample worksheets for maintaining cohort records (Current month and year is June 1976 for both.). 46

7 Using graphic presentation to depict directionality of trends 52

8 Graphic presentation of two subgroup trends for a measure 53

9 Time-series analysis of treatments for successive cohorts with identical times-in-treatment to compare different time-in-treatment trends 53

10 "True" cohort analysis tracing careers of four cohorts through treatment 54

11 An illustration of the lack of reliability of measures computed on small-sized groups 55

Tables

1 Alphabetized list of clients in a program during 1 month with their characteristics specified 39

2 Clients in table 1 subgrouped on the basis of treatment modality. 39

3 Clients in table 1 subgrouped first on the basis of treatment modality, then on the basis of sex. 39

4 Clients in table 1 subgrouped first on the basis of treatment modality, then on the basis of race-ethnicity 39

5 Information for computing "Percent drug-free (outpatient)" and value for each subgroup in tables 1-4 40

6 Constructing, maintaining, and dropping clients from cohorts in a program 45

7 Comparison of "normal" and "moving average" methods of constructing cohorts to meet minimum size requirements 50

PART 1

INTRODUCTION AND EXPLANATION OF SELF-EVALUATION

CHAPTER I

INTRODUCTION AND OVERVIEW

INTRODUCTION

This manual is designed to serve program administrators, counselors, and others interested in learning how to examine the effectiveness of a drug treatment program in achieving its goals. Initially, you should read the entire text to become familiar with the vocabulary and the order of presentation. The manual is meant to be used by persons with varying levels of skill in evaluation analysis--from the most elementary to the fairly sophisticated. As a result, persons with more advanced skills may not need to use some parts of the manual after initial familiarization.

The manual is partitioned roughly on the basis of familiarity with the evaluation process and skill in using evaluation techniques. This chapter discusses the rationale for evaluation, answers some questions often asked about evaluation, and discusses the role of program staff. Chapter II emphasizes the importance and functions of goals, objectives, and criterion measures in the evaluation process and describes the procedures for linking program goals with values obtained from measures of the clients' treatment outcomes. Potential problems, techniques for routinizing evaluation, and limitations of the self-evaluation process are discussed in chapter III.

Chapter IV provides an introduction to the measures and the methods for computing them. It is essential that you read and understand this section, particularly the differences between individual and group measures and the limitations and advantages of the measures. Chapter V contains instructions on how to compute the measures and brief discussions of what the measures may indicate about a program's functioning. Some limitations of each of the measures are also discussed.

Chapter VI will provide instructions on how to construct subgroups and cohorts. Chapter VII, the final chapter, will introduce trends and their importance in program evaluation.

Before you use the measures described in this manual for self-evaluation purposes, however, two suggestions are in order. First, you should read the rest of the chapters. Second, the worth of the measures in assessing your program is largely dependent on the quality of the information you receive. If the information you receive on clients does not reflect how the clients are actually doing, the usefulness of the measures for evaluating program functions is severely restricted. As the saying goes: "Garbage in, garbage out."

WHY BOTHER?

Accountability has become and will likely continue to be a critical concern of funding agencies. Delivery of human services is affected by accountability in at least two ways. First, drug treatment programs must comply with reporting requirements of Federal, State, and, sometimes, local funding agencies. Secondly, a program must divert some of its resources from the direct delivery of services to the provision of formal and objective justification of why the treatment program should continue to receive support. Hence, a treatment program must either engage an outside firm to conduct routine program evaluation or conduct its own self-evaluation.

However, not all programs have the necessary financial resources to conduct extensive, yet routine, evaluation efforts. This manual grew out of the need to provide treatment programs with some readily comprehensible guidelines for conducting self-evaluation efforts. The approach suggested is applicable for single programs or for an agency

composed of a variety of programs. In addition to proposing some general guidelines, several measures of treatment outcome are provided, along with instructions for computing the measures. As will be discussed below, the goals of a treatment program will determine the particular outcome measures most appropriate for the program. As a program administrator and/or the person in charge of program evaluation, you must keep in mind that everything in a general manual, such as this one, may not be suitable for your particular program. However, the suggested measures encompass the goals most generally adopted by different types of drug treatment programs, and it is anticipated that at least some of these measures will be relevant to most programs' goals.

You may ask: Why bother? There are several reasons. One goal of the evaluation process, perhaps the most important goal, is to enable you to continually upgrade the quality of patient care by learning more about what and how well your program is doing in regard to patient care. Moreover, your efforts may provide information that might otherwise not have come to your attention and may, therefore, help you in making better decisions. Evaluation efforts can often result in information useful for distribution to the public. Finally, the concept of accountability cannot be dismissed. Without continuing support of funding agencies, the most sincere counselor will not be able to do his or her job.

QUESTIONS AND ANSWERS

What kinds of questions do you want to answer? This manual is oriented to studying *program* effectiveness. While there is some discussion of studying clinical outcomes of a single patient or the patients of a counselor, the general concern of the manual is with your *program's* effects on all of your clients. No psychological tests are included and neither is there any in-depth treatment of goal attainment models. Again, this manual provides basic guidelines and steps for assessing information on an overall program.

Which staff members' help do you need? The answer is simple: *all of them.* Clinicians must be willing to report whatever information is needed, and other staff members must process it reliably. This does not mean, however, that an administrator will have to be constantly calling on them to fill out new forms. Most of the information required for the measures suggested in this monograph is *already being routinely collected.* The only extra requirement may be for someone to record information already routinely collected onto new worksheets so that the measures can be computed. Finally, if the information is kept on a month-to-month or other periodic basis, a program will be on its way to establishing its own data base.

How much will this self-evaluation cost? Since most of the data is already being collected, many of the guidelines suggested in this manual can be implemented at a minimum extra expense.

Will self-evaluation tell you everything you need to know? Simply put, no. Technically, a full-scale evaluation requires that unbiased research specialists (from outside your program) conduct an in-depth study of program operations and do followup on the clients. Such an elaborate study would include examining clinical procedures as well as clients' outcomes. Ideally, such a comprehensive effort would start even before a program was introduced into the community and would follow clients for several years after treatment. The study would not only get information on clients at intake, but would systematically take into account the nature and conditions of the community in which the program operates. Obviously, this is far beyond the scope of the suggestions in this manual. In fact, given the time and cost constraints under which most treatment programs operate, many important aspects of the evaluation process may have to be compromised. In particular, it will probably not be possible to use all the measures suggested here. Hopefully, though, each program will find in this manual some measures that are appropriate to its goals and objectives and that will help you know more than you knew before.

How do you know what to expect? Expectations will vary by type of program; age, race, and sex of clients; geographical location; and numerous other variables. However, caution is suggested in setting expectations based purely on intuition. For example, you might expect that, in order for your program to do well, 90 percent of the clients who were not employed at intake must be employed after 6 months in the program; however, intuition may result in unrealistic expectations if the local employment opportunities are poor during those 6 months.

One of the best ways to set up reasonable expectations is for you to compare your program with other similar programs. If you know of programs that are collecting information and computing measures that you use, you might consider exchanging information. Another way to know what to expect is to keep a record of your program's past performance so that you can compare it with current performance.

Later volumes in this monograph series will be devoted to providing outcome measures that have resulted from a national study and from the CODAP file. These volumes will include values for several types of measures, grouped according to general program types. Not all of the measures may be meaningful to a particular program. In a general sense, however, you may want to use them (at least initially) as data with which to compare the results of your own program efforts. As you will see, considerable effort has gone into constructing specific measures so that most programs can use some of the comparison data, if they choose.

Will this manual result in a cost-benefit analysis? Again, no. While some of the information collected could be used as a basis for this type of study, no instructions have been included for cost-benefit measures. As with the exclusion of detailed instructions on psychological scales, the goal is to keep the primary focus of the manual on program effectiveness, expressed in terms of clients' treatment outcomes.

Once you have done self-evaluation for a month or two, are you finished? That is up to you. It will certainly be more meaningful if you do not stop. Evaluation is usually an ongoing activity. It should be kept in mind that the information is useful not only to satisfy funding requirements, but can also be important in making decisions in areas that affect patient care as well as public relations with your community.

CLINICAL STAFF SUPPORT

An administrator must use his/her best powers of persuasion to gain the staff's cooperation. Like other kinds of research efforts, the quality of evaluation is often related to the amount of time and effort invested *and* to the commitment made to the evaluation project. Perhaps the most critical dimension of self-evaluation is the degree to which members of the clinical staff are committed to the endeavor. A spirit of cooperation is necessary in order to insure the reliability of reports based on clinical records. If the administrator has a person in charge of evaluation, that person must have the support of the program director. Furthermore, that person must be prepared to spend much of his/her time "selling" the clinical staff and keeping them informed of progress. Only in extreme cases should the person in charge of evaluation resort to using pressure from the director to obtain cooperation from a clinician.

If you use the results to avoid responsibility or if you use the data to serve as an excuse to remove a counselor from the program, then evaluation may be viewed as threatening and you stand a good chance of obtaining inaccurate information that may falsely make every counselor look good. Since this information is internal to your program, it is already subject to some bias. You must be careful in how you use the data. While it can help in making decisions, it cannot be used by itself to resolve all questions. Use evaluation as a streetlight to illuminate trouble spots, but do not lean on it for excuses when you must make decisions adverse to some of your staff.

The general disagreement between clinicians and administrators/evaluators about the way to spend time and money is not new. Often the clinician perceives that administrative and research staff exist to support the clinical effort. On the other hand, the administrative staff sometimes take the view that clinicians are invariably troublesome and fail to understand the practical realities of keeping a program alive. All too often underplayed is the reason for the program's existence in the first place--to serve a client and to meet a public need. As a result, the clinical and administrative staff must be interdependent if their mutual goals are to be met.

USEFUL RESULTS

Practical Information Needs

To use specialists in evaluation research and have them take charge of your evaluation needs can be quite expensive. Furthermore, some of the information they will generate may not really be very useful to you, or they might provide you with useful information in a form you do not understand. One advantage of routine self-evaluation is that you retain control. You control what aspects of program performance are to be looked at and you know just how everything is computed. Most importantly, you can evaluate your program with a minimum time and cost investment since you best know your program and its needs. Of course, the compromise is the objectivity and level of expertise that an outside person would otherwise provide.

The measures contained in part 2 of this monograph represent three general areas of outcome measures: *drug use, employment, and criminal activities.* Each of these general areas includes several different measures. You are free to choose the measures you want. It is strongly recommended, however, that you use at least one and preferably two measures

230-356 O - 77 - 2

from each area. In this approach to evaluation research, you are the judge of the most practical information needs.

Examples of Useful Results

Relying for the moment on your intuitive grasp of what a measure is and how one can be constructed, some examples of the results of using the instructions in this manual can be presented. Obviously, you may compute a value for a measure at any single point in time. That value will serve as an indicator of the status of your program's efforts at that time. The usefulness of a measure for evaluation becomes more obvious, however, when the actual value is compared to an "ideal" value which you have decided reflects satisfactory or successful performance or when the value is compared to the status of other similar programs elsewhere. Such comparisons help to give your program's value on the measure more meaning by providing standards or reference points for interpreting your results.

Beyond a single point in time, computing values for measures repeatedly can provide information useful for determining whether problems are developing, whether the needs of clients today are the same as those of former clients, whether treatment regimens are producing the expected results, why expected results sometimes do not occur, or whether attempted improvements in treatment are having the desired effects. For example, 6 months' values on the measure "Successful Completions as a Percentage of All Clients in Treatment" for a fictitious program are plotted in figure 1. Two pieces of information are immediately apparent: (1) about 20 percent of all clients normally complete treatment successfully each month and (2) for some reason, the gross successful completion rate for June was unusually low. Knowing the first fact helps reveal what the program is doing and what "normal" results can be expected so that we may decide whether program performance is satisfactory. Furthermore, since the value is fairly constant month-to-month, we are alerted to pay attention to unusual values which may signal problems, like the one for June. Then we know to investigate further to determine whether the problem is real (the low value may only have been a consequence of the fact that no clients were eligible for successful completion) and, if there really is a problem, to identify the cause(s).

For a program or agency manager, one persistent problem is deciding how to allocate limited funds and manpower. After programs are operating, one way to use instructions in this manual to help make that decision is to compare different parts of the program to determine how effective each is and where more effort seems to be needed. In an agency providing drug abuse treatment in several modality-environments, a first step might be to compare the modalities on a measure meaningful to all. Figure 2 portrays one such possible comparison on the measure "Percent Opiate-Free" between Therapeutic (residential) Community (TC), Outpatient Drug-Free (DFOP), Methadone Maintenance (MM), and Outpatient Detoxification (DETOX) modality-environments. At first glance, the TC modality-environment seems free of problems and quite effective. While the "Percent Opiate-Free" values are low for the DETOX modality, most of the clientele might always be new because of the shortness of the treatment term in this modality so that the low values may indicate that terminated clients are being quickly replaced. Such a possibility is supported by the constancy of the values and would reflect treatment effectiveness if confirmed by a short average length of time treatment slots remain open and by a high ratio of successful completions to all terminations. The interpretation is less clear for the other two modality-environments

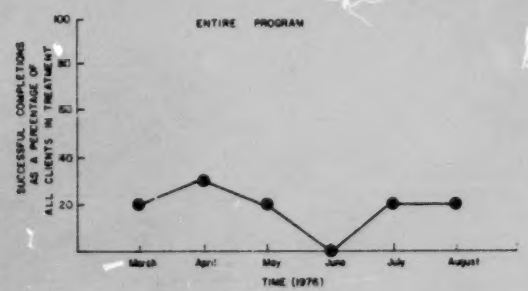

Figure 1. Determining normal treatment outcomes and identifying problems.

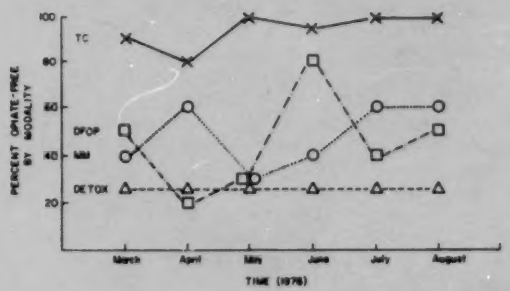

Figure 2. Comparing modality-environments to help determine resource allocation.

in figure 2, DFOP and MM. Additional factors must be considered, such as the proportion of DFOP clients who abuse nonopiates only, the ratio of new clients in both modality-environments to those who have been in treatment for several months, the availability of opiates in the community, and any changes in clientele or treatment efforts during the 6 months. Nevertheless, the comparison in figure 2 suggests initially that additional program resources may be less needed in the TC and DETOX modalities and provides some direction for investigating further which of the other two modalities is less effective and more in need of improvement.

Figure 3 provides a final example of the useful results that can be generated as a result of self-evaluation using the measures and instructions in this manual. The progress of groups of clients admitted to treatment during the same months (cohorts) can be followed in order, among other things, to gauge indirectly whether the quality of treatment remains high, to determine the optimum length of treatment, and to determine the effects of program changes at one point in time on clients at different stages of treatment. In the absence of any other information about the program, the clients, or the community, figure 3 is not particularly revealing, except to give the impression that the program is generally effective over time in curbing clients' opiate abuse. The low values for two of the cohorts in July should be investigated; perhaps a popular counselor left the program, perhaps the availability of opiates increased or the price went down, perhaps a change in treatment occurred, perhaps the staff spent an undue amount of time performing administrative rather than treatment duties, or all the "cleanest" clients may have simply left treatment during the same month. Whatever the reason(s), the process of regular self-evaluation provides results that help identify problems before they become unmanageable. Furthermore, figure 3 provides information that allows us immediately to discount the possibility that the two cohorts with low values are simply going through a bad phase of the treatment regimen since the timing is different for both and such a "slump" is not experienced at the same stage of treatment by all groups. This ability to eliminate some explanations for treatment outcomes so that time and resources will not be wasted is an equally valuable result of regular self-evaluation.

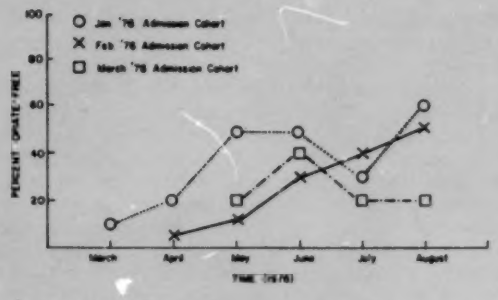

Figure 3. Comparing progress of groups of clients admitted to treatment during the same months (cohorts).

CHAPTER II
THE PROCESS OF EVALUATION

DEFINING GOALS AND OBJECTIVES

Many evaluation efforts fail because program goals have never been clearly and explicitly specified. In the absence of specific goals, it is difficult to construct and interpret measures of program effectiveness, and evaluation efforts may produce findings that are unrelated to the needs of a program.

What is a *goal?* A goal is the desired end product of your treatment efforts. Recent efforts at the Federal level have produced some measurable treatment goals which, to a greater or lesser degree, are common to all drug treatment efforts. These goals include:

1. Clients not using drugs in a manner that threatens personal safety or physical health, family or job adjustment;
2. Clients managing personal affairs in such a way that needs for food, clothing, shelter, transportation, and medical care are met in a responsible manner; and
3. Clients eliminating criminal activity, as measured by arrests and time incarcerated.

While these goals apply to most drug treatment programs, it is the responsibility of each program to add to this list additional goals uniquely suited to the program's treatment modality. For example, it is reasonable to expect methadone maintenance programs to specify goals different from those of a therapeutic community. Likewise, the meaning and relevance of employment goals will vary from program to program.

Goals must be established before the means by which a program intends to reach those goals are implemented. These ways, or paths, to attain goals are the *objectives* of a treatment program. In order to demonstrate the extent to which a program is moving toward (or failing to move toward) the achievement of a program's goals, measures must be designed which are indicative of these objectives. These *criterion measures* will then serve as the basis for self-evaluation.

Before measures (or indicators) of a program's goals and objectives are introduced, it is necessary to discuss priority setting. Some programs may not be able to achieve "success" on all their program goals simultaneously. Some goals may, by definition, limit the attainment of other goals. Or the program may have to limit certain efforts because of funding restraints. Consequently, it is necessary to list goals in order of priority. These priorities should reflect the importance of the various needs of clients relative to each other, the urgency of meeting clients' needs (e.g., long-term vs. immediate needs), and the availability of program resources. The consensus of the counseling staff should be integral to this process.

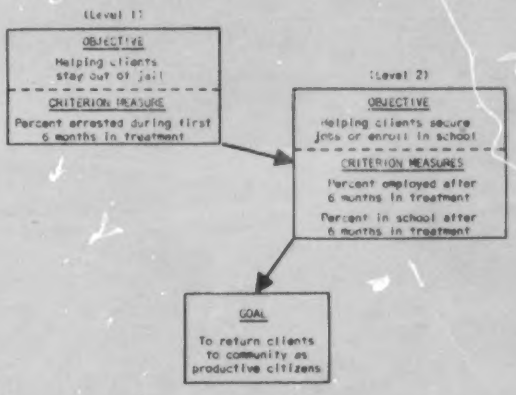
(Level 1)
OBJECTIVE
Helping clients stay out of jail
CRITERION MEASURE
Percent arrested during first 6 months in treatment
(Level 2)
OBJECTIVE
Helping clients secure jobs or enroll in school
CRITERION MEASURES
Percent employed after 6 months in treatment
Percent in school after 6 months in treatment
GOAL
To return clients to community as productive citizens

Figure 4. Relationship between goals, objectives, and criterion measures.

Further, the final priority ranking of specific goals will effectively define the program's emphasis. In addition to preparing the way for your self-evaluation, this process will discourage confusion of staff members, who might otherwise devote undue time and effort to accomplishing low-priority goals and thereby neglect higher-priority goals.

In pursuing program goals (e.g., the cessation of consequential drug use), numerous objectives will likely be met along the way. Objectives are reflections of the goals, and once objectives are established, specific outcome measures will become even more apparent. Also, just as priorities must be placed on different goals, it may be useful to specify the order in which objectives must be achieved.

An example of this strategy is depicted in figure 4. Level 1 objectives would be the most basic ends that are desired and, in this sense, would probably be the first tasks accomplished. Level 2 (and higher levels) would then follow, with each objective moving closer and closer to the desired goal. In figure 4, the ultimate goal is to return clients to the community as productive citizens. Toward this end clients must first be kept out of jail, an indicator of which is the percent arrested during the first 6 months in treatment. The next objective is to help clients secure jobs or enroll in some school. Indicators of this objective might be the percent employed and/or the percent in school after 6 months in the treatment program.

Recall that the general purpose of evaluation is to compare the effects of a treatment program with the goals it has set for itself and to use the results of evaluation in making decisions that will help a program better reach those goals. The effects of a program refer to the program outcomes. We use standards to compare the outcomes with the goals. Once initial goals are set, standards that will be acceptable for

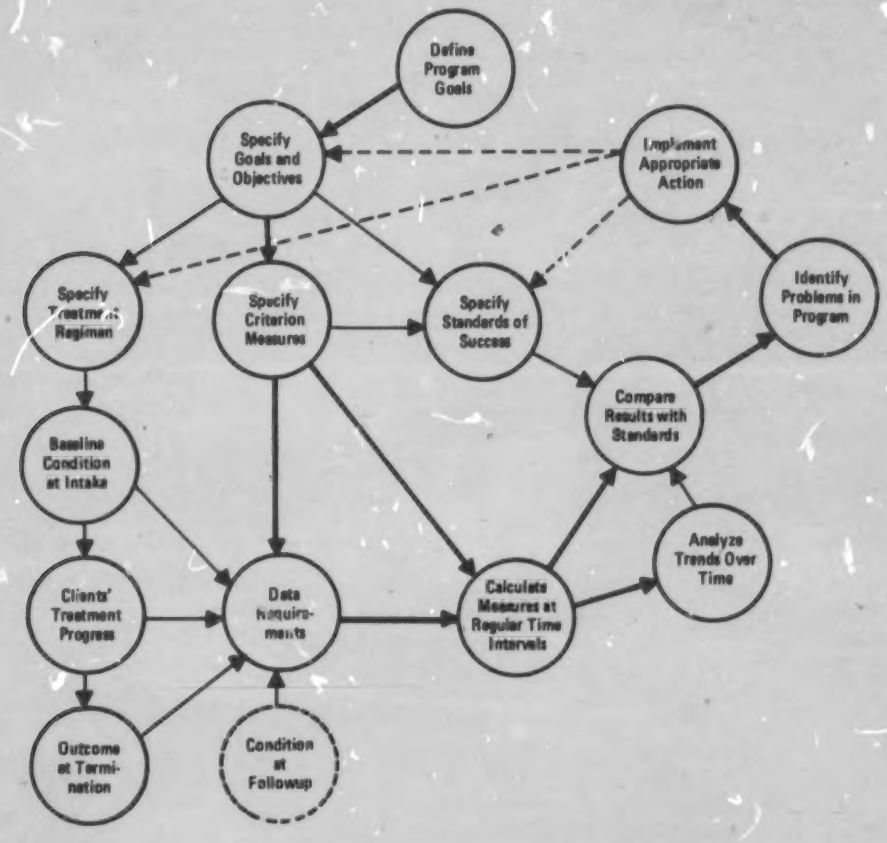

Figure 5. Summary of evaluation process.

definitions of success must be specified. Also, just as goals may be modified as you learn more about how the program operates, it may become necessary to change the standards of success. However, it is not acceptable to continually change criterion measures for any particular goal. Standards must reflect program goals, even changing ones, but in terms of the same criterion measures.

The particular standards chosen may depend almost totally on the particular program and its orientation(s). Whether you want to achieve a 10 percent or a 90 percent rate of employment largely depends on whether or not having a high rate of employment is an important program goal. For example, it would be most unusual to expect a high employment rate among residents in the first phase of a therapeutic community.

As has been mentioned, it may be desirable to set short- *and* long-term goals and have different standards for each. For example, a program may be willing to accept relatively high levels of unemployment until vocational rehabilitation personnel have been hired. Then a redefinition of program employment goals and a respecification of standards may be possible. Standards and goals should be consistent with one another.

Finally, you should take note that this monograph focuses on so-called "objective methods" of assessing program effectiveness. You should consider, if you have not already, getting feedback directly from your clients. In particular, you might want to develop a short questionnaire covering the important aspects of your program in addition to asking your clients if they felt they received quality care. Some programs even have an annual open meeting where clients and members of the community are invited to criticize the program openly and to ask questions.

SUMMARY OF THE EVALUATION PROCESS

Before proceeding further it will be useful to get a summary view of the evaluation process. Figure 5 presents an overview of that process. As discussed above, the first step is to specify and prioritize goals, objectives, and criteria measures. Next, the necessary baseline data must be obtained from clients at contact and/or intake. During treatment it will be necessary to collect information on the status of clients at regular, specified intervals. The length of time to be covered in each interval will be discussed in more detail later in the monograph. Information should be obtained on each client at termination. If the client is not available, then every effort must be made to obtain the information from the counselor(s), friends, acquaintances, or family.

Once information is obtained and recorded, the computational instructions in part 2 of this monograph become relevant. While no discussion of followup is included in this monograph, a logical next step would be to obtain additional information on clients after termination. Such information would allow you to assess longer-term consequences of treatment.

The last chapter in this volume deals with analyzing trends over time. This dimension of the evaluation process may prove to be one of the most meaningful, for it will allow a program to track changes through time.

CHAPTER III
IMPORTANT POINTS TO CONSIDER

KEEPING IT SIMPLE

Up to this point, we have been discussing the basic logic and potential worth of evaluation. Before getting into the outcome measures and computational instructions in part 2, however, there are some important guidelines that you should consider and keep in mind when doing evaluation. In fact, you may find it useful to refer to this chapter in the future to reorient yourself.

Part 2 includes some suggested measures that vary in their complexity and in the degree of time and effort required of you and your staff. If this is your first real involvement with evaluation, do not get overly eager. Start by selecting a single measure in which you have a particular interest. Try to get the information you need to compute the measure while involving as little of the staff's time as possible. Stop and think where you may already have some of the required information. Before actually collecting the information, draw up imaginary charts and practice computing the measure. Make sure you really understand what you are trying to do and that you are comfortable with the numbers you are going to have to use. Walk before you run!

As you practice with other measures, review how they differ from your first measure. If you find yourself getting lost in detail, retrace your steps. Go back to the original instructions and try to isolate what caused you to get confused.

BUILDING AND USING A DATA BASE

One orientation of this manual has been to keep the time and cost of your data collection to a minimum. However, data collection still requires some of your time. Consequently, you should try to get the very best use of the data by keeping in mind that, once collected, data can actually be used again and again. To do this, however, requires some planning and foresight on your part.

It is most important that you collect *baseline* data as soon as possible. Baseline data are usually the information you collect on your program's initial groups of clients or on persons at admission. Consequently, without such information, you will have no place from which to start. In general, you will eventually want to compare outcomes of groups of clients who have been in your program for the same amount of time. To do this, you will, in effect, want to see if their outcomes are different from their starting points. The technical guidelines for doing this will be discussed in part 3 along with a worksheet that you may want to use to organize your files.

It is strongly recommended that you set up a particular time period during each month to collect your information. You should then organize a set of files so that you will have current information readily available, but can also retrieve information from earlier time periods. As you will see later, this file will become the basis for monitoring changes (or *trends*) that may be taking place in your program. The worth of these kinds of observations will be discussed in more detail in part 3. Keep in mind that it will be necessary to go to some effort to get information, so try to maximize its worth by keeping it organized and readily available for future reference.

Two aims should be kept in mind. First, do not collect more data than you need. Secondly, always try to keep the number of new forms to a minimum so as to minimize staff burden. In this manual, almost all measures may be computed using data you routinely collect on CODAP forms. If you use only these measures, your most difficult task may

be to tighten up on the reporting of CODAP data. Again, if all your staff members are convinced that the data may have a positive effect on patient care, you will be more likely to get better cooperation.

THE REAL VS. THE IDEAL

Most attempts at evaluation cannot answer all the questions invariably asked by administrators, politicians, and the public. As you well know, statements on program effectiveness that are too general can have unanticipated consequences. Different interpretations and meanings may be read into the same findings. This may be particularly true of the findings of your evaluation efforts since you do not have the control of your patients that an experimenter working in an isolated laboratory may have. Consequently, you must always be somewhat cautious in making general statements about the findings of your evaluation efforts. Later in this manual you will be given examples of how different interpretations can be made of the same observation.

In addition to meeting external demands for accountability, the results of evaluation efforts can be used internally to further your general goal of effective patient care. From this perspective, results of evaluation should be used to confirm (or reject) your opinions about program aspects or practices that need improvement. It would *not* be prudent, however, to use your findings as the sole justification for a decision. Again, your efforts will necessarily be limited in terms of both a lack of control over important factors and, in all likelihood, the absence of followup information about long-term program effectiveness. In fact, information relevant to followup has been intentionally omitted from this monograph in order to keep the approach to self-evaluation at a level requiring minimal commitment of staff time and program funds.

The reason for all this caution is that to assess actual program effectiveness authoritatively requires stringent experimental research designs involving satisfying many more technical conditions than are necessary in order to use this monograph. We have deliberately not discussed issues such as random assignment or threats to internal validity.* While the ability to make general statements based on specific findings must therefore be restricted, acute concern with experimental research designs would take away from the primary goal of the manual--to help you get information that will assist you in the continual improvement of your program, and to do this in a straightforward manner.

*Some of the more complex issues and research designs are discussed in many of the references suggested at the end of this volume.

PART 2
SUGGESTED PROGRAM MEASURES AND COMPUTATIONAL INSTRUCTIONS

CHAPTER IV
INTRODUCTION TO THE MEASURES

PROGRAM MEASURES

Several program measures are provided in this manual, including group measures of (1) employment, (2) drug use, (3) criminality, (4) terminations, and (5) social stability. These measures were selected for the following reasons: their importance as indicators of program functioning, their importance to most drug programs, and the fact that information needed to compute them is currently being collected for the Client Oriented Data Acquisition Process (CODAP). The intention is to provide measures that are generally related to program goals and objectives, but which do not require the collection of much additional information. The first three types of measures can be computed directly from numbers you use to fill out the CODAP Client Flow Summary (CFS) and Activity Reports (ACR) each month. The termination rates can be computed from information on CODAP Admission (AR) and Discharge Reports (DR). Only the measures of social stability would require the collection of additional information.

Not all of the measures provided are appropriate to all programs, nor can all programs within a given agency be compared on these measures. *Programs, clinics, and agencies with different goals, objectives, and activities cannot legitimately be compared even on the same measures.* For example, employment during treatment is not usually an objective of residential treatment centers. This measure is therefore inappropriate for such facilities, and an employment rate for such facilities should not be compared to the employment rate for an outpatient methadone maintenance program. Consequently, you must decide which of the measures are appropriate for your program, based on your program's goals and objectives.

You will find instructions in chapter V on how to compute the measures as a *percentage* of eligible clients. The mechanics of computing percentages are simple, as illustrated below. First, instructions for each measure will define classifications of clients; for example, instructions for an employment measure might include:

E = number of clients employed (e.g., 25), and
T = total number of clients in the program (e.g., 50).

Then, a symbolic formula will be presented to illustrate the arithmetic involved in computing the measure (e.g., the employment rate):

$$\text{Employment Rate} = \frac{E}{T} \times 100 \ .$$

In words, this symbolic formula says:

1. Create the *fraction* representing "number of clients employed (E) *over* total number of clients in the program (T)," e.g.,

$$\frac{E}{T} = \frac{25}{50} \ .$$

2. Perform the division to get a *decimal*, e.g.,

$$50\overline{)25.0} = .5 \ .$$

3. Multiply the *decimal* by 100 to convert to a percentage, e.g.,

$$.5 \times 100 = 50\% \ .$$

We use percentages, rather than number of

clients, so you can compare the measures for your program at one time with the same measures at another time, even though the "base" number of clients in your program may be different. Also, values for measures in this manual have been computed using data from the CODAP system and from the Drug Abuse Reporting Program (DARP) and will be available to you later in two volumes. Expressing measures as percentages allows you to compare your program directly with these two sets of national values, even though the national values are based on many more clients than are treated by a single program.

As you will see from the following discussions about individual and group measures and about the importance of looking at the measures over time, the suggested measures have important limitations. Even with their limitations, however, the measures can tell you a great deal about how your program is doing and assist you in making some decisions. More importantly, they give you information on what areas of your program you should study in more depth.

Individual and Group Measures

Under current NIDA guidelines, one model available to you is a goal attainment treatment model. This means that counselors in consultation with individual clients set goals for each individual client to accomplish while in treatment. These goal attainment measures are calculated at the *individual* level and are essential for assessing how well any individual client is doing in the program. On the other hand, measures for individuals are difficult for program managers or counselors to summarize in order to see how well the *program as a whole* is doing.

Sometimes, the measures for all the individuals can simply be combined in some way to provide information about the program or about a subgroup in the program. But most measures that are ideal for evaluating individuals are not easily translated into indicators of program functioning, and many measures useful for program evaluation would be too burdensome if you first had to compute the measure for every individual. For example, if we define "employed" as "having any kind of job (including part-time)," then we can compute an employment measure for every individual client and assign a number to each client as his value for this measure, say 0 for unemployed and 1 for employed. The following example shows how we might "compute" an employment measure for individuals:

Client	Does client have any kind of job?	Value for Employment Measure
Bob	Yes	1
Jane	No	0
Daniel	No	0
Ruth	Yes	1
Carlos	No	0
Maria	No	0
Eldridge	Yes	1
Aretha	Yes	1

Notice that knowing a client's value (0 or 1) for the measure tells us nothing more about that client than we knew when we only knew the answer (yes or no) to the question, "Does the client have any kind of job?" Yet we must still transform the individual clients' values in order to indicate something about the employment picture for the program *as a unit* because the list of individual values makes no sense for a program--the questions we ask about the program are different from those we ask about the individuals.

One way to get an indication about the program is to use the information portrayed in the example above to compute a *program* employment measure, based on the example formula discussed on page 11. By counting the number of employed clients (E = 4) and the total number of clients (T = 8), we can fill in the "employment rate" formula:

$$\text{Employment Rate} = \frac{E}{T} \times 100 ,$$

$$= \frac{4}{8} \times 100 ,$$

$$= 50\% .$$

Now, we know that half of all the clients in the program are employed (and that half the clients are unemployed), and we can begin to say something about the way the program is functioning. For example, if all clients are out of work at admission, then it appears that being in the program improves the employment situation for some of them. On the other hand, if one of the program's goals is for all clients to get and keep a job, then the fact that the program's employment rate is only 50 percent is an indication that the program's goal is not being met.

The difference between measures on individuals

and measures on groups such as a program is crucial--it reflects the fact that we expect different things from groups than we do from individuals, and we therefore ask different questions. Hence, we need different answers. In the example above, the 1 and 0 (and yes and no) values for individuals make no sense when used for a program (unless the clients in the program are either all unemployed or all employed). Likewise, the 50 percent value for the program makes no sense when used for an individual (a client either has a job or he does not--part-time employment is defined as "a job" and thus no client can have 50 percent of a job). Thus, while the measures to be presented in this part of the manual tell you nothing about any single individual client, they do tell you how well the *group* of clients is doing.

Discrete and Continuous Measures

The operation of a drug treatment program is a continuous process, and the best measures of its functioning are those which reflect that continuity of operation. Unfortunately, such *continuous* measures are so costly and time consuming that client treatment would have to be neglected in order to collect the information required to compute the measures and to conduct the evaluation.

Alternatively, the time and effort required for evaluation can be kept at a more reasonable level by using measures based on the situation at a single point in time or which summarize all the events occurring between two points in time. We call such measures *discrete* because they artificially divide a program's continuous operation into separate, isolated incidents. Some information and precision is thus lost and some limitations have to be placed on interpretation of the measures, all of which are justified because the evaluation task becomes manageable.

Examples of one type of discrete measure are the items of information you report each month on your CODAP Client Flow Summary (CFS) and Activity Reports (ACR): e.g., number of clients in treatment, number of clients employed, number using other drugs, etc. Specifically, CODAP CFS Item 19, "Client Census - Last Day of Month," is a discrete measure in that it portrays the program's status only on the last day of the month. That is, the measure does not reflect any changes that may have occurred *during* the month, such as admissions, discharges, or reassignments to different environments and modalities. Hence, the monthly client census measure must be interpreted cautiously, since there is no guarantee that it accurately portrays what is going on in the program all, or even most, of the time. On the other hand, reporting a client census measure for *every* day of the month would be too burdensome. Hence, the more manageable discrete measure is used on the assumption that the situation on the last day of the month should not be very different from the situation on any other day.

Discrete measures can also be computed by summarizing events occurring between two points in time, say between the last day of the previous month and the last day of the current month. The employment rate *at one point in time* discussed above would be computed:

$$\text{Employment rate} = \frac{\text{Number of clients employed on last day of current month}}{\text{Total number of clients in treatment on last day of current month}} \times 100.$$

The same measure based on a summary of employment events and treatment events *during the month* could be computed:

$$\text{Employment rate} = \frac{\text{Number of clients employed at any time during the month}}{\text{Total number of clients in treatment at any time during the month}} \times 100.$$

For such a measure, any client who held any job (full- or part-time) for any time (even if only 1 day) during the month would be included once (and only once) in the numerator, and any client (including admissions and discharges during the month) in treatment for at least one day of the month would be counted *once* in the denominator. Since this type of measure is based on a summary of events during a time period, it more closely approximates the continuity of program operation and can be interpreted as describing a program's typical situation during the month, rather than a program's (possibly untypical) situation on a particular day.

Both the measures just discussed are discrete. Perhaps the best way to summarize the difference between continuous and discrete measures is to use an example based on film. The continuous nature of program operation is similar to viewing a motion picture or movie. The measures we present

in this manual for use in evaluation are like snapshots, taken either at one point in time or by repeatedly exposing the same frame of film several times. In either case, we cannot completely depict the *process* of program functioning; rather, we can only "freeze" the action. Like a snapshot, discrete measures are accurate--they portray the situation that exists at the time--but they cannot describe for us the complete process of which the situation is a part. Nevertheless, discrete measures, like snapshots, are affordable, simple, quick, and informative when interpreted with care.

THE ELEMENT OF TIME

Time is particularly important for evaluating treatment outcome and program operation. Treatment outcome depends heavily on the length of time the program has to work with the client. The time period for which information is collected to compute evaluation measures can affect both the values and the usefulness of the measures. And comparison of measures repeatedly computed through time is far more informative than single values of the measures computed at one point in time.

Report Period

Throughout this manual, we will use a calendar month as the standard time period covered by the measures. We recommend that you also collect information about your program, compute the measures, and evaluate them on a monthly basis for several reasons. First, the CODAP system already requires monthly reports, and this makes some of the information needed for some of the measures readily available. Second, a month provides enough time so that unusual events will not have an undue influence on the measures. Third, collecting information and computing measures on a monthly basis is not too demanding. Finally, monthly evaluation provides the frequent looks at your program needed to identify troubles before they become serious problems, to assess the impact of program changes you institute, and to compare current functioning with past performance.

Time in Treatment

Generally, the longer a client remains in a program, the better are his chances of benefiting from treatment. *Cohort* enumeration is a way of organizing clients into groups based on the length of time they have been in the program. A cohort is defined as "an aggregate of individuals . . . who experienced the same event [in this manual, "admission to treatment"] within the same time interval."* Later discussions will provide instructions on creating cohorts (see chapter VI) and on analyzing measures computed for cohorts (see chapter VII) in order to evaluate your program's effects over time.

Trends: Program Functioning Through Time

Knowing the values of a program's measures for the current month is important--only in this way can we determine the present situation in the program. Then, we can begin to conduct evaluation by comparing values of the measures with goals we have set for the program. In addition, we can compare the program's current status with identical measures computed for similar programs elsewhere.†

But evaluation is not complete without comparing present program functioning with the program's past performance. Analysis of trends enables us to make this crucial comparison. A *trend* is the general direction of movement *over the course of time*. Looking at the current status of the program relative to past values of the measures can suggest (but cannot prove) whether program functioning is changing and the rate at which any changes are occurring. Unfortunately, the history of a program's values on a measure--i.e., a trend--cannot reveal the *reasons* for stability or change in a program's effectiveness. But by repeating and comparing measures over time, we can have more confidence in the results of the measures at any single point in time, and we can use another criterion--change or stability of the measures--for evaluating program effectiveness. Because trends are so important to evaluation, an entire chapter of this manual has been devoted to the analysis of measures over time.

---

*Ryder, Norman B. "The Cohort As a Concept in the Study of Social Change." *American Sociological Review*, Volume 30 (December 1965), pp. 843-861.

†Two volumes will be available to you for comparing your program's values on most of the measures in this manual with values for the same measures computed using data from programs in the CODAP system and in the Drug Abuse Reporting Program (DARP).

CHAPTER V
COMPUTATION OF THE MEASURES

INTRODUCTION

This chapter will provide instructions on how to compute the measures. In addition to providing computational instructions, this chapter will also give some interpretations of what the measures might mean and some limitations on their use. However, it is essential that users of the measures read chapter IV and understand the nature of discrete measures and the basic difference between measures for individuals and group measures.

Instructions have been provided for measures related to the following areas: (1) employment, (2) drug use, (3) criminality, (4) termination rates, and (5) indicators of social stability.

Two Ways to Compute Some of the Measures

Two formulas are provided for computing the employment, drug use, and criminality rates. Computation (the arithmetic) is equally easy for either formula.

The first formula for each measure is based on all clients in treatment at any time during a reporting period, regardless of whether or not they were in treatment at the end of the time period. In order to use the first formula, information must be collected on all clients discharged or admitted during the report period as well as on clients in treatment for the whole period. You cannot, therefore, compute the first formula for these measures directly from information on the CODAP Client Flow Summary (CFS) and Activity Reports (ACR). However, if you use the first formula for each of the measures, you will be able to compute the kind of discrete measures that summarize events between two points in time (see page 13). Because summarizing events during a time interval gives a better picture of a program or subgroup than do discrete measures based only on the situation on a single day, this first formula may be worth the slight additional effort and time needed to collect the required information.

The first formula for the employment, drug use, and criminality measures will be labeled as a "DARP" formula. Measures computed for your program or for a subgroup using the DARP formulas can be directly compared with values based on programs in the Drug Abuse Reporting Program (DARP) for the same measures. A separate volume containing the DARP values on selected measures will be available to you through NIDA.

For those who wish to compute *program* measures directly from the CODAP monthly report forms, the second formula for computing each of the measures is provided. For most measures, *the second formula is based only on clients in treatment on the last day of the reporting period. Persons discharged prior to the last day of a reporting period are not counted anywhere in this type of formula.* Since the second formula for each measure is related to the CODAP reporting forms, the second formula will be labeled a "CODAP" formula. The advantage of using the CODAP formula is that measures *for the total program* can be computed directly from information on the monthly CODAP reports, but the results must be interpreted more carefully when only 1 day's operation is represented (see pages 13-14).

Notice that the discussion of the CODAP formulas mentions only *program* measures. Information on the monthly CODAP reporting forms concerns *all* clients in treatment at the end of the month. Many numbers you report on the form make no distinctions among the clients in terms of their individual time in treatment, drug usage pattern,

age, sex, race and ethnicity, or even treatment modality. If your clientele is significantly varied on these or other characteristics, there may be good reason to believe that the values on measures of treatment outcome will be different for subgroupings based on these characteristics. For example, the unemployment rate for young black males is higher than for any other age/race/sex subgrouping in the general population, indicating that young black males have special needs apart from any drug abuse problems in finding and keeping jobs. Computing your program's unemployment rate using the figures from your monthly CODAP CFS and ACR reports, you could neither determine how different the employment situation is for the young black male clients in your program nor how effective your treatment and employment efforts are in meeting that subgroup's special needs.

The fact that measures can be computed only for the program as a whole when you use the information reported on the CODAP CFS and ACR forms is a limitation of the *figures* you insert into the CODAP formula, not of the formula itself. As a matter of fact, any CODAP formula can be used equally as well as the corresponding DARP formula for computing subgroup and cohort measures (subject to the more limited interpretation discussed above), *if* the characteristics needed for classifying each client into a subgroup and cohort are first linked to the client's status on the items to be used in the formulas. Indeed, another separate volume of values from the CODAP system for some of the measures provides values for subgroups and cohorts by first classifying clients into subgroups and cohorts on the basis of information from the CODAP Admission Report (AR), and then computing the measures using the CODAP formulas in this manual and information on each client's CODAP Discharge Report (DR). Since the volume of values from the CODAP system will be regularly updated and published by NIDA, we offer two recommendations: (1) become familiar with the CODAP formulas by computing measures for your total program using the information from your monthly CODAP CFS and ACR reports, and (2) set up and maintain a system that will allow you to classify clients into and compute measures for subgroups and cohorts using the CODAP formulas (see chapter VI and appendix A). Then you can evaluate your program's effects on subgroups of clients and over time (i.e., with cohorts), as well as its effects on *all* clients, relative to the treatment outcomes being reported across the nation.

Some General Instructions

Each measure introduced below will be accompanied by one or more formulas, definitions for the terms used in the formulas, and a brief discussion of limitations and possible interpretations of the measure. Some general instructions and rules apply to all the measures, however, and are discussed here.

Techniques for computing the measures are portrayed symbolically, with letters and arithmetic formulas. The actual arithmetic involved in computing percentages has already been discussed (see pages 11-12). A detailed explanation of the symbols, definitions, and formula will be presented for the first formula of the employment measures below: "Percent Employed Full Time." Instructions for all measures thereafter will present the symbols, definitions, and formulas without repeating the explanation. *It is important therefore, that you review the instructions for computing percentages (pages 11-12) and that you read the entire section of instructions for the measure "Percent Employed Full Time."*

Letters are used throughout this manual in all the formulas as a kind of "shorthand" for the way clients are classified before computing the measures. Then the formulas symbolically define the arithmetic to be performed. You should not be alarmed by all these symbols since they actually represent only categories of clients and simple arithmetic, not "statistical jargon." The definitions for the letters will furnish the instructions you need to classify clients appropriately. Once you have determined the numbers associated with the letters and carefully substituted the numbers for the corresponding letters in the formulas, simply perform the indicated arithmetic operations in the order: Add and subtract first; then multiply and divide.

Every client *eligible* for a measure should be counted at least once: in the denominator (bottom term) of the formula [usually "total number of clients in treatment" (N)]. Clients are *eligible* for being counted only if they meet the requirements for the measure. *In general, a client can meet the requirements for a measure only if, during the reporting period, he could have experienced the treatment outcome being measured.* For example, you may keep a client in a treatment slot during the period although he spent the entire month in jail on a traffic

charge. But when you compute a drug use measure for the month, that client is *not eligible* for being included anywhere in the drug use measure formula because he could not have experienced the outcome event: using drugs.*

Not all of the clients eligible for a measure will be included in the *numerator* (top term) of the formula for that measure. For example, when we compute the measure "Gross Termination Rate," clients still in treatment should not be included in the numerator.

For measures with several categories of clients in the numerators of the formulas, it is possible for a single client to meet the requirements of the definitions for two or more categories. *However, do not count a client as being in more than one category in the numerator.* For example, the following instructions might appear for an employment measure:

E = number of clients employed full- or part-time.

H = number of full-time housewives *not* employed and *not* in school.

S = number of full-time students *not* employed.

T = total number of clients eligible for employment and in treatment during the month.

$$\text{Employment rate} = \frac{E + H + S}{T} \times 100 \quad .$$

As you can tell from the formula, housewives and students are treated as if they are employed. But you may have housewives and students who also really are employed. Such clients should be counted only once in the numerator. In other words, count an employed student either as a student (S) or as employed (E) but not both. This is an important rule to keep in mind for *all* the measures: *do not include clients in more than one category of the numerator or of the denominator.*

When a CODAP formula for a measure is presented, the letters used as symbols for the categories of clients will have a "prime" (i.e., E', H', T', etc.) to distinguish them from similar but not identically defined categories in the corresponding DARP formula. Also, the definitions of the categories will be followed by the initials of the CODAP form and the item number where the appropriate numbers for the formula can be located (e.g., "ACR Item 18"). However, you cannot use the information on some of the CODAP forms to make comparisons *between* subgroups of clients on the various measures; you can only compute measures on *all* your clients (see pages 15-16).

EMPLOYMENT MEASURES

A major goal of many outpatient programs is that all clients be self-supporting through legitimate jobs. Of course, most programs never have *all* the clients employed at one time because clients are usually in various stages of treatment. But to get an *indication* of how close to the goal a program is, three measures of the employment situation are provided below: (1) the percent employed full time, (2) the percent employed part time, and (3) the percent unemployed.

The employment situation is subject to many influences, some of which are beyond the control of program personnel. Nonprogram factors include the availability of jobs in the community, discrimination against drug users or former drug users, the difficulty in getting jobs for persons with poor previous employment records or with arrest records, and level of clients' job skills at admission to treatment. Program-related factors include staff skills in job training and placement, referral patterns between treatment agencies, and the emphasis placed on employment by the staff.

You should attempt to determine what factors, both program related *and* those unrelated to the program's functioning, are influencing the values of your program's employment measures. Identification of the factors influencing the employment measures could provide important information on what your program should or can do if you want to increase the employment rate among your clients. For example, if your program's "Percent Employed Full Time" is low and you find that this is due to a large number of admissions during the month of clients with poor job skills, you may want to increase the emphasis your program places on vocational education and job training. On the other hand, a low "Percent Employed Full Time" measure may not be influenced by *any* change in the program if the entire community is experiencing an economic recession accompanied by widespread job layoffs.

*Disregard for purposes of this example the possibility that the client may continue to have access to drugs while in jail. In any case, the category, "total number of clients in treatment" (N), for all the measures *should* include only clients *who* actually *could have received treatment* during the reporting period, whether the period is a month or just a day.

Two formulas are presented for each of the employment measures. The first formula (DARP) for each measure summarizes the employment status of clients *during the entire month* and requires information on unemployed clients who are either housewives or students. Information for the second formula (CODAP) can be found on the monthly CODAP Client Flow Summary (CFS) and Activity Reports (ACR) if you want to compute the measures only for your entire program. Whenever you use the CODAP formulas, whether for the entire program or for subgroups in the program, assign clients who were in treatment during the month to categories on the basis of their employment status *at any time during the month*, count housewives and students not otherwise employed as unemployed, and then use in the formula only clients *in treatment on the last day of the month.*

For all the formulas, a client, to qualify as *employed*, must have earnings subject to income taxes from a legitimate job, including self-employment (see *CODAP Handbook*, AR Item 13, for definition of "Employment").

Percent Employed Full Time (Full-Time Employment Ratio)

This measure provides the simplest and most direct indication of how close a program (or a subgroup in the program) is to meeting the goal of full-time employment for all clients. Its major limitations are that it does not enable us to identify the specific factors that affect the employment situation (e.g., we cannot separate the effects the program is having on employment from the effects caused by the local job market) and it does not count persons with part-time jobs as employed.

DARP Formula

The measure of full-time employment computed using the DARP formula counts as employed housewives and students who do not have any kind of paying jobs, thus not penalizing a program for clients who would not be employed even if they had no drug abuse problem. More information must be collected to use this formula than to use the CODAP formula, however, and the information on the CODAP forms cannot be used because it does not match the definitions of formula terms. In addition, the full-time employment status of the program (or a subgroup) will be *overstated* somewhat because clients can be counted as employed full time if they held a job for as little as 1 full working day--there is no requirement for this DARP formula that a client be gainfully employed in the same job every working day of the entire reporting period.

E_f = number of clients employed full time *at any time* during the reporting period.

H = number of clients who were full-time housewives, *not* employed and *not* in school during the entire time period.

S = number of clients who were full-time students, *not* employed, during the entire time period.

N = total number of clients in treatment *at any time* during the time period, including clients discharged or admitted during the period.

(DARP) Percent Employed Full Time

$$= \frac{E_f + H + S}{N} \times 100 .$$

The formula says that the number of persons employed full time as a percentage of all clients in treatment during a time period is equal to the sum of the number of clients employed full time plus the number of housewives plus the number of students, divided by the total number of clients in treatment *at any time* during the time period, multiplied by 100. The symbolic formula instructs you to perform the following operations:

1. *Add* the number of clients employed full time (E_f), the number of clients who were full-time housewives (H), and the number of clients who were full-time students (S).
2. *Divide* the sum ($E_f + H + S$) from (1) above by the total number of clients in treatment (N) to get a decimal number.
3. *Multiply* the decimal number from (2) above by 100 to get the percentage of all clients who were employed full time at any time during the period.

CODAP Formula

This formula allows use of information from the CODAP CFS and ACR forms for the program's measure (but not a subgroup's) of full-time employment. The problem of overstating the full-time employment situation, mentioned above for the DARP formula, is *not* eliminated since this CODAP formula also reflects employment status *at any time* during the month, but *only for clients in treatment on the last day of the month.* On the other hand, program effectiveness in reaching its full employment goal *may* be *understated* if the clients who terminate treatment before the last day of the month are also the ones who have been able to get and hold jobs as a result of the program's efforts.

E'_f = number of clients *in treatment on the last day of the month* who were employed full time *at any time* during the month (ACR Item 7 line 2).

N' = total number of clients *in treatment on the last day of the month* (CFS Item 18).

(CODAP) Percent Employed Full Time

$$= \frac{E'_f}{N'} \times 100 \ .$$

The formula says that the number of persons in treatment on the last day of the month and who were employed full time at any time during the month as a percentage of all clients in treatment on the last day of the month is equal to the number of clients employed full time at any time during the month and still in treatment at the end of the month divided by the total number of clients in treatment on the last day of the month.

Percent Employed Part Time (Part-Time Employment Ratio)

This measure is computed similarly to the "Percent Employed Full Time" measure above. Its interpretation is still limited by the fact that we cannot separate the effects the program is having on part-time employment from the effects of other factors which, in the case of part-time employment, are more complex than for full-time employment. It is *not* a limitation that clients employed full-time (and full-time housewives and students) are not included in the categories for this measure since our interest is in getting an indication of the *part-time* employment situation *only*. In fact, it is recommended that you compute this measure any time you compute the "Percent Employed Full Time" measure in order to help interpret the total employment situation for the program.

DARP Formula

Since we have already counted full-time housewives and students not employed in the DARP formula for "Percent Employed Full Time," they are not counted as part-time employees here.

E_p = number of clients employed part time *at any time* during the reporting period, *excluding* clients employed full time at any time during the period.

N = total number of clients in treatment at any time during the time period, including clients discharged or admitted during the period.

(DARP) Percent Employed Part Time

$$= \frac{E_p}{N} \times 100 \ .$$

CODAP Formula

E'_p = number of clients *in treatment on the last day of the month* who were employed part time *at any time* during the month *excluding* clients employed full time at any time during the period (ACR Item 7 line 1).

N' = total number of clients *in treatment on the last day of the month* (CFS Item 18).

(CODAP) Percent Employed Part Time

$$= \frac{E'_p}{N'} \times 100 \ .$$

Percent Unemployed (Unemployment Ratio)

While the two previous employment measures indicate how close the goal of full employment is, the "Percent Unemployed" measure provides an indication of the size of the job still to be done. More importantly, in conjunction with computing the "Percent Unemployed" for subgroups and cohorts, the measure provides information about what the program needs to do since the clients with the employment problems can be directly identified. For example, if your program's "Percent Unemployed" is fairly low (say, 15 percent) but the value on the same measure is extremely high (say, 60 percent) for the subgroup of young black males, program efforts should focus heavily (although not exclusively) on this group's special needs, particularly on job training and efforts with employers to overcome discriminatory hiring practices. On the other hand, a

230-358 O - 77 - 5

program with even a moderate "Percent Unemployed" (say, 30 percent) would not necessarily need to make any changes if, at the same time, almost all of the clients out of work are in the cohort of latest admissions, which would indicate that the employment situation improves as treatment progresses.

Because a measure of unemployment* can be used to provide information directly relevant for remedying the problem, we have included formulas for such a measure. In addition, these formulas have been used to compute values for programs in the CODAP and DARP systems so that you may compare this aspect of your employment situation with the experience of other similar drug treatment programs nationwide. For these reasons, you will probably find the "Percent Unemployed" measured the most useful of the three employment measures.

---

*The employment measures presented to this point correspond to versions of the "Labor Force Participation Rate" computed by the U.S. Bureau of Labor Statistics (BLS). It will become apparent that this "Percent Unemployed" measure (both DARP and CODAP) is simply 100 minus the full- and part-time "Labor Force Participation Rate." Such an unemployment measure is not the same as the BLS "Unemployment Rate" so often discussed in the media and available locally from Departments of Employment Security. But there are two compelling reasons for your using the "Percent Unemployed" measure presented here:

1. The required data are regularly available, at least for the CODAP formula, while data for the BLS Unemployment Rate must be generated.
2. The BLS Unemployment Rate is based primarily on persons unemployed and seeking work. But for a drug treatment program, the possibility that some clients may not be looking for work is perhaps more important than determining that jobs are not available--clients making no effort to find employment are probably the ones most in need of immediate, intensive attention by program staff.

On the other hand, the trend of a program's "Percent Unemployed" measure over time should be compared with the local Unemployment Rate trend, even though the two measures do not have the same meaning. Regardless of program efforts, the community job market will make attaining the program's employment goals more difficult during general economic slumps and comparing the two measures will help explain higher "Percent Unemployed" values during such periods.

This measure must be interpreted as carefully as the previous two employment measures since the same mix of program related and nonprogram factors influence the unemployment situation. Furthermore, the definition of *unemployed* is quite narrow, in that it counts clients as unemployed only if they held *no* job at any time during the reporting period. As a result, *the unemployment situation is likely to be understated* by both the formulas below.

DARP Formula

We have already counted in the DARP formula for "Percent Employed Full Time" housewives and students who are not employed because they are productively occupied full time and not seeking work. Thus, even though full-time housewives and students do not earn taxable income from a job, they are *not* to be counted as unemployed here.

U = number of clients *unemployed during the entire reporting period.*

N = total number of clients in treatment *at any time* during the time period, including clients discharged or admitted during the period.

(DARP) Percent Unemployed

$$= \frac{U}{N} \times 100 .$$

CODAP Formula

$E'_f + E'_p$ = number of clients *in treatment on the last day of the month* who were *employed* full or part time *at any time* during the month (ACR Item 7, sum of both lines).

N' = total number of clients *in treatment on the last day of the month* (CFS Item 18).

(CODAP) Percent Unemployed

$$= \frac{N' - (E'_f + E'_p)}{N'} \times 100 .$$

Relationships Between the Employment Measures

It should be noted that all the terms from all the DARP employment formulas are related in the following manner:

$$N = E_f + E_p + H + S + U .$$

This equation means that no client should be

counted as being in more than one of the categories on the right side of this equation, regardless of the employment measure being computed. In addition, the three (DARP) employment formulas are related in the following manner:

$$\text{(DARP)}\quad \left(\frac{E_f + H + S}{N} \times 100\right) + \left(\frac{E_p}{N} \times 100\right) + \left(\frac{U}{N} \times 100\right) = 100\% \ .$$

This equation means that the value of any one of the DARP employment measures can be computed by subtracting the sum of the values of the other two DARP employment measures from 100. Also, an "overall" (full-time and part-time) employment ratio can be determined, either by subtracting from 100 the value of the "Percent Unemployed" measure computed using the DARP formula, or by adding together the values of the "Percent Employed Full Time" (DARP) and "Percent Employed Part Time" (DARP) measures. Hence, it is not necessary to compute all three employment measures from scratch; you can determine the value of any third measure from the values of the other two employment measures.

Notice also that there is no separate term in any of the CODAP employment formulas for unemployment. The number of clients unemployed for the entire month and still in treatment at the end of the month (U') can be determined from the numerator of the formula for "Percent Unemployed" (CODAP):

$$U' = N' - (E'_f + E'_p) \ .$$

This equation means that the same rules discussed above for the DARP employment terms and formulas also apply to the CODAP terms and formulas. Thus, the CODAP employment formulas are also related:

$$\text{(CODAP)}\quad \left(\frac{E'_f}{N'} \times 100\right) + \left(\frac{E'_p}{N'} \times 100\right) + \left(\frac{N' - (E'_f + E'_p)}{N'} \times 100\right) = 100\% \ .$$

MEASURES OF DRUG USE

Four drug use measures are provided below: (1) the percent opiate-free, (2) the percent using nonopiate drugs, (3) the percent of outpatients free of all drugs, and (4) the percent abusing alcohol. No single one of these measures is sufficient by itself to demonstrate program effectiveness in meeting its goals concerning drug use. All the measures are needed to begin to get an overall picture of the complexity of drug use; even with all four of these measures available, only *indications* of program functioning can be obtained.

As with employment, many factors influence drug use, some of which are not related to clients' experiences with the drug treatment program. The availability and cost of drugs, the commitment of clients to overcoming their drug abuse problems, and the types of drugs clients used and the quantity and frequency of use before treatment are some of the nonprogram factors which influence drug use. Program related factors include the program's philosophy and counselors' attitudes toward use of different kinds of drugs, substitution of other drugs for the drug of abuse, the amount of control exercised over clients by the program, and the model or combination of models of treatment used.

Again, only by trying to identify both the program related and nonprogram factors that influence drug use can you determine the information needed to use these measures as aids in decisionmaking about treatment efforts. For example, a fairly high "Percent Opiate-Free" value is, by itself, difficult to interpret without more information. On the one hand, if you know that the latest admissions are the only clients using opiates, the high value for all clients would indicate that your program is relatively successful in helping curb opiate abuse. On the other hand, such a high "Percent Opiate-Free" value may provide little information about program functioning if law enforcement efforts have drastically restricted the supply of opiates during the reporting period.

Some of the categories used in the three nonalcohol drug use measures are defined below in terms of detecting drug use by urinalysis. Depending on whether or not your program uses such tests, here are some instructions for how your method for detecting drug use must be used in the formulas. If you do *not* conduct urinalyses, counselors must determine by any possible method (e.g., clients' self-reports, observation, etc.) which clients used drugs during the reporting period and the kinds of drugs used. All the clients must then be categorized so that the measures may be computed on the basis of your counselors' reports, rather than on the basis of urinalysis reports. The values for the drug use measures will then be comparable with each other and with the values for the programs that are reported in the separate volume based on the DARP system.

If urinalyses *are* conducted in your program,

it does not matter in computation of the measures whether such tests are made for all clients or for only a sample of clients. If *all* clients are tested, categorize the clients and compute the measures in the usual manner, according to the definitions. If urinalyses are conducted for only a *sample* of clients, *count only the clients actually tested* for drug use in your computation of the measures. While the results may vary slightly from the total program's (or subgroup's) actual values on the drug use measures, the presumption is made that urine tests for a properly chosen sample* of clients are representative of tests for all the clients. Thus, measures based on a properly chosen sample should accurately represent the program's values.

For *all* the drug use measures, including the measure "Percent Abusing Alcohol," the numbers used in the formulas are defined as numbers of clients, *not* numbers of urine tests, positive urines, or binge drinking episodes. Thus, a client who has used at least one kind of the generic drug at least once (has at least one positive urine) is counted only once in the "user" category for the measure, regardless of how many other similar drugs he also used, how frequently he may have used the drug, or how often he may have been checked for use (by either the counselor or by urinalysis). Furthermore, count the "number of *clients* reported on" in the denominators (bottom terms) of the formulas, *not* the total number of checks by counselors or tests conducted.

The three measures for nonalcohol drug use are intended *primarily* to give indications of *illicit* drug use. For this reason, definitions of "drug use" categories printed below *exclude* users of *only* prescribed and licitly obtained drugs. Use of prescribed and legally obtained Methadone or other opiate antagonist would not, therefore, be considered opiate use for the measure "Percent Opiate-Free." Use of other medications prescribed for legitimate medical reasons would similarly not be sufficient to classify a client in the user category for any of the measures. Without exception, this exclusion of clients using *only* licit and prescribed drugs from the user categories applies in the DARP formulas below; as a result, collecting *all* the information required for the DARP formulas will involve some additional effort to identify these clients. The same exclusion of users of only prescribed drugs generally applies for the CODAP formulas, with the exception that some complications are introduced when the measures "Percent Using Nonopiates" and "Percent Drug-Free" are computed for the entire program by using figures from the monthly CODAP CFS and ACR forms. These additional conditions will be discussed with the instructions for the CODAP formulas below.

If you are concerned with licit prescription drug use, we recommend that a separate measure be created, independent of the measures defined below. To do so, count the clients who *only* use licit prescribed drugs, divide the number of clients in this category by the total number of clients, then multiply by 100 to compute a "Percent Using Licit Prescribed Drugs Only" value.

Percent Opiate-Free
(Opiate-Free Ratio)

This measure provides information directly relevant to evaluating program effectiveness in meeting the goal of eliminating the use of nonprescription and illegally obtained opiates. It has three major limitations: (1) specific factors affecting illicit opiate use are not identified (e.g., when the value is high, the measure does not tell us if it is high because the program has been effective or because the supply has been reduced); (2) *all* opiate use and, thus, program efforts to reduce opiate *dependency* are *not* reflected; (3) the value may *overstate* program success if some of the clients in treatment (and counted in the denominators of the formulas) have *never* used opiates (i.e., the program gets credit for treating a problem that some clients never had).

A related measure, the "Percent *Using* (Illicit) Opiates," may be computed by subtracting

---

*A "properly chosen sample" can be derived by any of several methods. It is assumed here that the large programs which need to undertake urine screening on a sample basis are also aware of and correctly use these sampling methodologies to determine which clients' urines will be screened. Coincidentally, the sampling for urine screening will support inferences for the entire program about program measures computed on the sample clients only.

Unfortunately, sampling often reduces the numbers of clients in subgroups below acceptable levels (see page 39). As a result, subgroup detail will be restricted when computing drug use measures based on urine screening samples of clients. For a helpful discussion of sampling techniques, see W. Mendenhall, L. Ott, and R. L. Scheaffer, *Elementary Survey Sampling*, Wadsworth Publishing Co., Belmont, California, 1971.

from 100 the value resulting from either of the two formulas below.

DARP Formula

O = number of clients reported to have used one or more nonprescription opiates at least once during the reporting period. (Use all sources of information available, including urinalyses.)

N_d = total number of clients in treatment *at any time* during the period, including clients discharged or admitted during the period, for whom information about use or nonuse of drugs during the period is available. (Use all sources of information available, including urinalyses.)

(DARP) Percent Opiate-Free

$$= \frac{N_d - O}{N_d} \times 100 .$$

CODAP Formula

O' = number of clients *in treatment on the last day of the month* who each had at least one positive urine during the month for nonprescription opiates (ACR Item 10.c line 1).

N'_d = total number of clients *in treatment on the last day of the month* who each received urinalysis tests one or more times during the report month (ACR Item 10.a).

(CODAP) Percent Opiate-Free

$$= \frac{N'_d - O'}{N'_d} \times 100 .$$

Percent Using Nonopiates
(Nonopiate Drug Use Ratio)

This measure provides a fairly direct indication of the use of nonprescription nonopiate drugs in a program (or subgroup). Values on this measure cannot be related to values of the other drug use measures in this manual like the employment measures were (see pages 20-21) since every client using both nonprescription opiates and nonopiates is counted in two measures, resulting in more "users" than there are clients in the program if all the drug measures are combined.* Nevertheless, this measure is suggested because it can contribute important information about program efforts in dealing with nonopiate drug abuse. The measure is *not* limited by the fact that clients using either prescribed or nonprescribed opiates *only* are not counted in the "user" category since we are interested in use of nonprescription drugs besides opiates. Interpretation of the measure is limited, however, by the fact that we cannot separate effects of the program from effects of nonprogram factors on the drug use situation and we cannot identify, using this measure, the nonopiate drugs most abused. The overall drug abuse situation with which a program must deal is somewhat *understated* by this measure because clients using only opiates or only prescribed nonopiates are not counted.

DARP Formula

Clients using multiple nonopiate drugs are counted only once in this formula and then only if one of the drugs is obtained without prescription. As a result, the value obtained by using this formula even further *understates* the overall drug abuse situation.

D_{no} = number of clients reported to have used one or more nonprescription nonopiate drugs at least once during the reporting period. (Use all sources of information available.)

N_d = total number of clients in treatment *at any time* during the period, including clients discharged or admitted during the period, for whom information about use or nonuse of drugs during the period is available. (Use all sources of information available.)

(DARP) Percent Using Nonopiates

$$= \frac{D_{no}}{N_d} \times 100 .$$

CODAP Formula

This formula for the measure "Percent Using Nonopiates" is subject to some serious restrictions and is provided as much to illustrate the difficulties that can arise when data intended for one purpose are used for another as it is to be used for evaluation. The source of the difficulty is the fact that, unlike the DARP formula for this measure, a client using *multiple* nonopiate drugs *may* be counted more than once if positive urine test results are reported for all the nonopiate drugs used. A client taking both a depressant and a stimulant may have reported

*Also see the discussion of a similar complication (i.e., the possibility of "double counting") in the instructions for the CODAP formula for this measure.

two positive urinalyses, one for each category of drugs. As a result, that same client may be counted once in ACR Item 10.c line 2 and once again in ACR Item 10.c line 3, as indicated by the following definitions and formula.

Stim' + Dep' = number of clients *in treatment on the last day of the month* who each had at least one positive urine during the month for nonprescription stimulants (Stim') *and/or* at least one positive urine during the month for nonprescription depressants (Dep') (ACR Item 10.c line 2 plus ACR Item 10.c line 3).

N'_d = total number of clients *in treatment on the last day of the month* who each received urinalysis tests one or more times during the report month (ACR Item 10.d).

(CODAP) Percent Using Nonopiates

$$= \frac{\text{Stim}' + \text{Dep}'}{N'_d} \times 100 .$$

The client with positive urines for both categories of drugs will be counted twice in the numerator of the formula but only once in the denominator. Such "double counting" restricts the usefulness of the measure in several ways:

1. Overall drug *use* is understated in the measure because all prescription drugs and nonprescription opiates are ignored and because the categories "Stimulants" and "Depressants" may not include all drugs being abused. While double counting may serve to offset this understatement somewhat, the procedure is not arithmetically legitimate and so cannot be called a true "correction."
2. Theoretically, at least, double counting can result in the measure "Percent Using Nonopiates" (CODAP) having a value greater than 100 percent. This possibility does not limit the usefulness of the measure for self-evaluation, which is essentially the process of comparison with standards, as long as the standards (e.g., your program's past values, values for programs nationwide) are computed in exactly the same manner or are independent of the method of calculation (e.g., a goal of *zero* percent using nonopiates). It may be embarrassing, however, to report publicly that more than 100 percent of your clients are using nonprescription nonopiates, especially to an audience unaware of or unsympathetic to the difficulties of data collection.
3. For most other measures in this manual, any value has one relatively unambiguous interpretation: The number of clients affected by the characteristic being measured affect the calculated percentage once; all other clients do not affect the percentage value. But double counting results in a value whose meaning is unclear: An unknown number of clients may be affecting the percentage value more than once. Hence, to improve treatment effectiveness, you are uncertain whether to focus your efforts on a certain percentage of your clients, all of whom have the same minimal problem, or on a smaller and unknown number of clients whose problem has been "counted" as more serious. Interpretation of measures of program effectiveness is difficult enough without introducing more complications.

Since the primary purpose of the measures in this manual is to provide means for self-evaluation (i.e., comparison), the formula above is included for programs to use in comparing present with past performance and with drug use goals. Because of the ambiguities in interpretation, however, values on this measure for programs nationwide will not be included in the third volume of the series.*

---

*An alternative calculation for the measure "Percent Using Nonopiates" (CODAP) can be performed using the following definitions and formula:

P'_u = number of clients *in treatment on the last day of the month* who each had at least one positive urine during the month *for any drug* (ACR Item 10.b).

O' = number of clients *in treatment on the last day of the month* who each had at least one positive urine during the month *for nonprescription opiates* (ACR Item 10.c line 1).

N'_d = total number of clients *in treatment on the last day of the month* who each received urinalysis tests one or more times during the report month (ACR Item 10.a).

(CODAP) Percent Using Nonopiates (Alternate)

$$= \frac{P'_u - O'}{N'_d} \times 100 .$$

Percent Drug-Free Outpatients (Outpatient Drug-Free Ratio)

This measure provides an indication of the effects of treatment on eliminating abuse of any drugs. Because the differences between the DARP and CODAP formulas for this measure are more extreme than for most of the other measures, possible interpretations and limitations are discussed separately for each of the formulas. Regardless of these differences, however, a value calculated from either the DARP or the CODAP formula for the "Percent Drug-Free Outpatients" measure does not provide information permitting separation of program from nonprogram factors affecting the drug use situation.

DARP Formula

Consistent with previously discussed drug use measures, this formula counts as drug users those clients taking only nonprescribed and illicitly obtained drugs. Hence, the name of the measure may be misinterpreted because clients using *only* prescribed and legally obtained drugs are considered "drug-free." As a result, values obtained using this DARP formula are likely to *overstate* the degree to which actual use of *all* drugs has been reduced. On the other hand, the "Percent Drug-Free Outpatients" measure based on this DARP formula is particularly suited for programs which attempt to curb drug abuse by substituting legitimately prescribed drugs (e.g., maintenance modalities) since evidence of the treatment drug use is ignored.

It should be noted that the title of the measure mentions only *outpatient* clients because some treatment environments seek to control the client's access to drugs. Theoretically, many hospital inpatients and residents of therapeutic communities cannot obtain drugs except licitly by prescription. For such programs, this measure would be superfluous, always having a value of 100 percent. Evidence exists (see the volume for programs' values on this measure nationwide based on DARP) to suggest that *all* programs, not just those treating outpatients only, may find it useful for self-evaluation to compute their actual drug-free ratio.

D_i = number of (outpatient) clients reported to have used one or more *illicit* drugs at least once during the reporting period. (Use all sources of information available.)

N_d = total number of (outpatient) clients in treatment *at any time* during the period, including clients discharged or admitted during the period, for whom information about use or nonuse of drugs during the period is available. (Use all sources of information available.)

(DARP) Percent Drug-Free (Outpatient)

$$= \frac{N_d - D_i}{N_d} \times 100 .$$

CODAP Formula

Unlike the DARP formula above and unlike other measures of drug use, the category of drug use for this formula makes no distinction between prescribed or nonprescribed drugs, licitly or illicitly obtained. Hence, the value computed with this formula provides a direct indication of the actual proportion of clients who use *no* drugs of *any* kind; i.e., clients who are actually "drug-free." For this reason, use of this formula by programs whose treatment regimen includes prescription of substitute drugs for the drug(s) of abuse will result in a significant *understatement* of program effectiveness. Furthermore, this formula can be computed only for the program as a whole when only information from the CODAP CFS and ACR forms is used. Programs with a goal that clients become truly "drug-free" will find this measure most useful, but caution must be used in interpreting high, but not 100 percent, values: clients using only drugs prescribed for legitimate medical problems unrelated to their drug abuse problems may be counted as "users" in this CODAP formula, thereby resulting in a possible (and small) *understatement* of program effectiveness.

P'_u = number of clients *in treatment on the last day of the month* who each had at least one positive urine during the month *for any drug* (ACR Item 10.b).

Unfortunately, this alternative calculation may be less useful because the term "P'_u" includes users of *prescription* drugs (both opiate and nonopiate). If, as in some treatment models, substitution of a prescribed and licitly obtained drug for the drug(s) of abuse is a legitimate solution of clients' problems, this alternative calculation turns treatment efforts into part of the problem. If elimination of *all* drug use is the goal, the "Percent Drug-Free Outpatients" (CODAP) measure below is a more direct indicator of treatment effectiveness and is easier to calculate.

N'_d = total number of clients *in treatment on the last day of the month* who each received urinalysis tests one or more times during the report month (ACR Item 10.a).

(CODAP) Percent Drug-Free

$$= \frac{N'_d - P'_i}{N'_d} \times 100 .$$

Percent Abusing Alcohol
(Alcohol Abuse Ratio)

This measure is provided because alcohol consumption, particularly for some age groups and in some subcultures, is related to the use of certain nonopiate drugs and may complicate the problem of nonalcohol drug abuse and its treatment. Since alcohol may be legally obtained almost everywhere nationwide, definitions below are not given in terms of licit or illicit use. However, the definitions of alcohol abuse do differ for the DARP and CODAP formulas and will be discussed separately. Nevertheless, for both formulas, factors related to treatment efforts are not separated from nonprogram factors in the values computed. Furthermore, since little is known about the causes of alcohol abuse and since alcohol is so easily obtainable, the task of interpreting this measure in terms of program effectiveness in treating abuse of all drugs is difficult. Perhaps the best general interpretation of an alcohol abuse ratio is that the clients affected have problems of which drug and alcohol abuse are only symptoms so that treatment efforts must continue beyond merely stopping drug abuse.

DARP Formula

The definition of alcohol abuse for this formula is regular consumption of the equivalent of 8 ounces or more of 80-proof liquor. By "regular consumption" is considered "average daily consumption," so that over the course of a week, for example, the total amount of alcohol consumed equals the equivalent of 56 ounces or more of 80-proof liquor. Equal quantities may not necessarily be consumed on a daily basis, and usually not all 56 ounces are consumed during a single day. The equivalent of 8 ounces of 80-proof liquor is one of the following:

1. Eight shot glasses or mixed drinks of 80-proof liquor,
2. One half-pint of 80-proof liquor,
3. One fifth of wine, or
4. One six-pack of 16-ounce cans of beer.

While other definitions of alcohol abuse are legitimate, the particular one used here has widespread acceptance, is consistent with NIAAA guidelines, and has been used in computing values for programs in the volume based on DARP. If you choose to use this definition, some extra effort will be required of your counselors to determine clients' quantity and frequency of alcohol consumption. Finally, the measure of alcohol abuse based on this formula and definition may either *overstate* or *understate* the problem because clients have different capacities for absorbing alcohol; i.e., clients who can consume the indicated quantity regularly without it interfering with normal functioning will be counted as "abusers" while the clients whose well-being is severely impaired after consuming smaller quantities less frequently will be ignored.

AA = number of clients who abused alcohol at any time during the reporting period. (Use all sources of information available.)

N_a = total number of clients in treatment *at any time* during the period, including clients discharged or admitted during the period, for whom information about use or nonuse of alcohol during the period is available. (Use all sources of information available.)

(DARP) Percent Abusing Alcohol

$$= \frac{AA}{N_a} \times 100 .$$

CODAP Formula

Alcohol abuse, for this formula, is defined as use of alcohol to an extent sufficient to interfere with a client's physical, mental, or social well-being (see *CODAP Handbook*, ACR Item 11). Notice that this definition does not specify a minimum quantity or frequency of alcohol use. Values computed using this formula can reflect the different effects alcohol use has on different clients, but the reliability of the values may suffer since classification of clients depends on the subjective judgment of your counselors whose ideas of "interference with well-being" may differ. For a program's values, information from the monthly CODAP ACR and CFS forms can be used in the formula.

AA' = number of clients *in treatment on the last day of the month* who each had an alcohol abuse problem at any time during the month (ACR Item 11).

N' = total number of clients *in treatment on the last day of the month* (CFS Item 18).

(CODAP) Percent Abusing Alcohol

$$= \frac{AA'}{N'} \times 100 .$$

CRIMINALITY MEASURES

One major aspect of the problem of drug abuse is the fact that use of most drugs is often illicit. Since there may be some connection between illicit drug use and other illegal activities undertaken to finance drug purchases, two measures of criminality are presented: (1) percent arrested and charged, and (2) percent engaged in illegal activities. In conjunction with efforts to reduce drug usage, many programs have as a major goal helping clients to replace with legal and productive activities any illegal activities which place them in jeopardy of involvement with the criminal justice system.

Criminality, like employment and drug use, is influenced by many factors. Criminal experience prior to treatment, the legality of the drug(s) abused before treatment, the effectiveness of law enforcement efforts, the differential exposure of clients to opportunities to engage in crime, and the prevailing attitudes in the community toward crime and drugs are factors over which the program has little control. Program related factors include the program's philosophy and counselors' attitudes toward clients with criminal records, program efforts to substitute legal and productive activities for illegal ones, the amount of control exercised over clients by the program, and success in direct treatment of the drug problem.

Perhaps the most important consideration in interpreting these criminality measures is whether or not the values are related to clients' drug abuse. If the clients arrested are charged with drug offenses or with offenses undertaken specifically to finance drug purchases, a major part of the remedial action should be intensified efforts to curb the drug use. If, on the other hand, the offenses are less directly related to the clients' drug abuse, such as robbery by unemployed clients to get money to eat or crimes against persons within the family, program efforts may need to be directed over several areas, such as job training and placement, curbing drug use, and counseling in interpersonal relations.

Percent Arrested and Charged
(Arrest and Charge Ratio)

This measure is a direct indicator of the encounters clients have with the law enforcement and criminal justice systems. It provides information about the most serious effects of criminal behavior from the client's drug abuse perspective: possible interruption of his treatment. The measure has two important limitations: (1) it does not enable us to identify specific factors that affect criminality, and (2) it depends so heavily on the level of law enforcement efforts that values for single time periods may not be comparable (although an average value over several time periods would be useful).

DARP Formula

For this formula, the definition of clients arrested and charged does *not* include those charged for disorderly conduct, vagrancy, or misdemeanor traffic violations. You or your counselors will have to determine the offense for each client so that the minor offenders can be excluded. The "arrested and charged" category *does* include clients charged with the following types of crimes: (1) gambling, running numbers, etc.; (2) prostitution or pimping; (3) stealing, forging, etc.; (4) drug violations; (5) crimes against persons; and (6) all other felony crimes. Occasionally (e.g., twice a year), you may find it useful to calculate the percent arrested and charged for each of the six categories of crime.

C = number of clients who each were arrested *and* charged at least once during the reporting period, excluding clients charged only with disorderly conduct, vagrancy, or misdemeanor traffic violations.

N = total number of clients in treatment *at any time* during the period, including clients discharged or admitted during the period.

(DARP) Percent Arrested and Charged

$$= \frac{C}{N} \times 100 .$$

CODAP Formula

For this formula, no distinction is made between types of offenses. Thus, clients

330-356 O - 77 - 3

arrested and charged for *any* offense are included in the category "arrested and charged." While this formula can be computed using information from monthly CODAP ACR and CFS forms, clients arrested and charged during a month are more likely to be incarcerated and, thus, not in treatment on the last day of the month so that values computed using this formula may significantly *understate* the extent of criminality in your program.

C' = number of clients *in treatment on the last day of the month* who each were arrested and charged for *any* offense at least once during the month (ACR Item 12).

N' = total number of clients *in treatment on the last day of the month* (CFS Item 18).

(CODAP) Percent Arrested and Charged

$$= \frac{C'}{N'} \times 100 .$$

Percent Engaged in Illegal Activities (Illegal Activities Ratio)

This measure may be a better indicator of client criminal activities than the one just discussed because the "Percent Arrested and Charged" measure is influenced so much by the level of police activity. However, this measure cannot be computed from information on any of the CODAP forms, it is not related to the criminality measure in the volume of values for programs based on DARP, and you or your counselors must collect and organize some information you may not ordinarily gather. On the other hand, this measure may tell you more about your program's effectiveness in dealing with actual client criminality since the "Percent Arrested and Charged" is only an index of "discovered" crime. Even so, this measure is limited by the possibility of reporting error and by clients' reluctance to reveal, even to counselors confidentially, their illegal activities; thus, values may *understate* the situation.

Since you have to collect additional information in order to compute this measure, it is suggested that the measure be computed on all clients in treatment at any time during the time period, including clients discharged or admitted during the period. And, while it is unlikely that clients will report disorderly conduct, vagrancy, and traffic violations as "illegal activities," it is suggested that any such reports be *excluded* from the category of clients engaged in illegal activities.

I = number of clients reported as having engaged in one or more illegal activities during the reporting period. (Use all sources of information available.)

N_i = total number of clients in treatment *at any time* during the period, including clients discharged or admitted during the period, for whom information about illegal activities during the period is available. (Use all sources of information available.)

Percent Engaged in Illegal Activities

$$= \frac{I}{N_i} \times 100 .$$

TERMINATION MEASURES

While employment, drug use, and criminality measures are important indicators of how a program is doing, measures concerned with terminations can also be used to evaluate programs and to identify strengths and weaknesses on the basis of the reasons why clients leave. Ideally, the goal is to have all clients successfully complete treatment (whether or not each client stays in treatment for the full term formally specified in a treatment regimen). A program has met this goal if, for each client terminating, it can be determined that the client would not be substantially served by remaining any longer in treatment. Of course, most programs never are able to make this claim about *all* clients terminating at any one time because clients leave treatment for a variety of reasons. But to get an indication of how close to the goal a program is, two types of termination measures are provided: (1) gross termination ratio, and (2) termination ratios by type of discharge. In addition, some relationships between the formulas for these measures are discussed, similar to the discussion of employment measures (see pages 20-21). Using these measures, you can generate information on what proportion of your clients is terminated during a time period and gain some insight about the most important reasons for leaving.

All the termination measures in this section can be computed either from information on the CODAP Client Flow Summary (CFS) form or on the Client Discharge Report (DR) form. As a result, each letter used as a symbol for a category will have a prime (e.g., T'), all the definitions of categories will be followed by the initials of the CODAP form and item number where the information for the formula can be located, and all the formulas will be labeled "CODAP." In addition, the same formulas were used to compute values on these

measures for programs in the DARP system. As a result, an extra formula is not presented for each measure, and the values you compute using the formulas printed below can be compared directly to the values in the volume based on DARP programs.

It should be noted that the categories in these formulas differ somewhat from earlier categories for which information was taken from the monthly CODAP forms. Many of the categories below require that you *combine* entries from the forms to include all clients in treatment *at any time* during the report month. In any formula involving clients discharged *during* the month, the base category of *eligible* clients (i.e., the denominator) must include *all* the clients who, during the same time period, could have also experienced the outcome (e.g., termination for all reasons, successful completion of treatment, etc.). For this reason, the formulas below are slightly more complex than previous CODAP formulas, but you gain the ability to summarize events during the entire month and to compare results from a single computation with a DARP value as well as with the CODAP value in subsequent volumes.

Furthermore, "late" AR and DR reports are included in the computations for the month in which they are reported, rather than for the actual month of admission or discharge. Including late reports in the current month's computations will result in values slightly off the mark for the current month, but it is justified in terms of the time and effort otherwise required to match late reports with their month of occurrence and then recompute values for the earlier months. Also, over a period of several months, the slight errors in assignment can be expected to have little effect.

Percent Terminating for All Reasons (Gross Termination Ratio)

The "Gross Termination Ratio" tells you what proportion of all clients in treatment during a time period were discharged, regardless of the reason for leaving. This measure may be computed for the entire program directly from information on your monthly CODAP Client Flow Summary Report (CFS).

As its name implies, the measure provides no details about reasons clients left treatment, whether program related or not. And you must interpret the value each month on the basis of the situation in your program for that particular month. Unlike most of the previous measures, *both* high *and* low values on this measure can indicate program effectiveness, depending on the treatment status of clients. For example, if your program is designed to provide 1 year of treatment and half your clients successfully complete a year of treatment during the same month, a "Gross Termination Ratio" of 57 percent for that month is commendable. On the other hand, a value of about 50 percent *every* month reflects the fact that many clients are leaving before the end of the prescribed term, indicating either that the treatment plan should be revised or that the program is having difficulty retaining clients long enough for treatment to be effective.

Of all the termination measures presented in this manual, this is the only one related directly to *total* clientele; for this reason, this measure should regularly be computed. (But also see the measure "Adjusted Termination Ratio" below.)

T' = number of Discharge Reports submitted at the end of the month (CFS Item 16 plus CFS Item 17).

N'_t = total number of clients in treatment *at any time* (i.e., eligible for termination) during the month (CFS Item 13 plus CFS Item 14 plus CFS Item 15).

(CODAP) Percent Terminating For All Reasons

$$= \frac{T'}{N'_t} \times 100 .$$

Percent Terminating by Reason for Leaving (Reason-Specific Termination Ratios)

Termination information is most useful for evaluation purposes when the reasons for clients' leaving treatment are compared. For example, evaluation of a program would be less favorable if most of the terminations result from clients leaving against staff advice than if most of the clients successfully completed treatment. You can directly compare the different reasons for termination and their importance for your program relative to each other by distinguishing the terminated clients on the basis of their reasons for leaving. As discussed previously (see page 28), expressing the number of clients leaving for each reason, either as a percentage of all clients in treatment or as a percentage of all terminations during a time period, will provide greater flexibility in making comparisons. Since the minimum use you will want to make of these measures is to compare them with each other, it is suggested that most "Reason-Specific Termination Ratios" be expressed as percentages of *all terminations* during a month, because such a method directly relates the effects of each reason for

discharge to the other types of terminations. (But see the measure "Adjusted Termination Ratio" below.)

In order to compute the "Reason-Specific Termination Ratios," the reasons for termination have to be taken from the CODAP Discharge Reports (DR) submitted each month with the Client Flow Summary. On the CODAP Discharge Report is an item labeled "Reason for Discharge" (Item 6). That item lists eight possible reasons for a client's termination. To determine the number of clients who left for each of the eight reasons, separate the Discharge Reports and group them according to the code numbers entered in DR Item 6. Then count the number of reports in each group.

Several, though not all the possible "Reason-Specific Termination Ratio" formulas are presented below. The first three formulas will probably be most useful for self-evaluation in that the relationships revealed by comparing these measures with each other can provide some direct insights into program strengths and weaknesses. As indicators of program effectiveness, the remaining measures are more ambiguous but can be useful if interpreted with caution.

Percent Terminating Adjusted for Successful Completions (Adjusted Termination Ratio)

The "Gross Termination Ratio" presented earlier disregards the distinction between those terminations that are indicative of program success and those which may reflect program ineffectiveness. Since successful completion presumably is a desirable outcome of treatment, it would be useful to have a measure which does not penalize the program for these successful completions, yet provides information about the proportion of the total clientele whose treatment has been unsuccessful so that corrective action may be taken. This measure attempts to accomplish such a purpose by focusing on the proportion of "problematic" terminations.

S'_t = number of Discharge Reports submitted at the end of the month for clients considered to have successfully completed treatment (DR Item 6 code 1).

T' = number of Discharge Reports submitted at the end of the month (CFS Item 16 plus CFS Item 17).

N'_t = total number of clients in treatment *at any time* (i.e., eligible for termination) during the month (CFS Item 13 plus CFS Item 14 plus CFS Item 15).

(CODAP) Percent Terminating Adjusted for Successful Completions

$$= \frac{T' - S'_t}{N'_t} \times 100 .$$

For public relations purposes, it may be useful to have a measure of program "success," reflecting the proportion of *total* clientele successfully completing treatment during a time period. Such a measure, the "Successful Completion Ratio," can be easily computed by subtracting the value of the "Adjusted Termination Ratio" from the value of the "Gross Termination Ratio."

Successful Completions as a Percentage of All Terminations

One condition of successful completion is that the client has been in treatment for some specified length of time (or a reasonable approximation of the term called for by the formal treatment regimen). However, for any one month it is probable that the majority of clients will have been in treatment for a lesser period and are thus *ineligible* for termination by reason of successfully completing treatment. To try to express program effectiveness based on *all* clients in treatment, using a measure like the "Successful Completion Ratio" above, will result in a low value that *understates* program success merely because most of the clients could not have successfully completed treatment, regardless of program efforts.

By expressing the number of successful completions as a percentage of the number of terminations, rather than of *all* clients, the problem of understating program effectiveness due to ineligibility can be avoided* while providing an indication of the extent to which clients successfully complete treatment relative to other reasons for termination. High values on the measure presented here are a direct indication of program effectiveness. Low values indirectly reflect ineffective program efforts, even during the months when very few clients are eligible for successful completion, because they indicate that most

*It should be noted that the problem is only avoided, not *solved*. Only by taking into account the different lengths of time clients have been in treatment can the problems associated with eligibility for successful completion be eliminated entirely. See the discussion below of cohorts for a technique that corrects for the effects of time in treatment.

clients are leaving treatment before they even become eligible for successful completion.

S'_t = number of Discharge Reports submitted at the end of the month for clients considered to have successfully completed treatment (DR Item 6 code 1).

T' = total number of Discharge Reports submitted at the end of the month (CFS Item 16 plus CFS Item 17).

(CODAP) Successful Completions as a Percentage of All Terminations

$$= \frac{S'_t}{T'} \times 100 .$$

Left Voluntarily as a Percentage of All Terminations

Of the measures presented in this manual, this one probably comes closest to the meaning of the common term, "Split Rate." For the purpose of program improvement in order to better serve persons with drug abuse problems, this measure is extremely useful. Low values indicate at least that a program is not alienating clients so much that they quit treatment and probably reflect clients' judgments that the treatment is beneficial. A high value, on the other hand, is a direct indication that, at best, clients are leaving treatment before the program has a chance to extend full benefits and, at worst, that the program itself is doing something to drive clients away.

V'_t = number of Discharge Reports submitted at the end of the month for clients classified as having voluntarily left before completing treatment (DR Item 6 code 6).

T' = total number of Discharge Reports submitted at the end of the month (CFS Item 16 plus CFS Item 17).

(CODAP) Clients Leaving Voluntarily as a Percentage of all Terminations

$$= \frac{V'_t}{T'} \times 100 .$$

Rule-Breakers as a Percentage of All Terminations

A client may also leave treatment as a result of a "program decision to discharge client for noncompliance to program rules" (DR Item 6 code 5). While rules and discipline are always necessary, with the result that some clients will have to be discharged for breaking the rules, it may be useful for self-evaluation to compute a measure of disciplinary discharges. Such a measure, particularly if its values are high or become high over time, may indicate that treatment is not being responsive to the needs of clients or, worse, that program efforts are being focused more on enforcing rules than on helping clients.

RB'_t = number of Discharge Reports submitted at the end of the month for clients discharged by program decision for noncompliance with program rules (DR Item 6 code 5).

T' = total number of Discharge Reports submitted at the end of the month (CFS Item 16 plus CFS Item 17).

(CODAP) Rule-Breakers as a Percentage of all Terminations

$$= \frac{RB'_t}{T'} \times 100 .$$

Transfers as a Percentage of All Terminations

Termination of clients as a result of transfer or referral to another program or agency can be presumed to be beneficial for the clients, but, without additional information, it is more difficult to interpret terminations of this type in terms of program goals. Normally, however, programs make more efficient use of resources if client intake is selective, designed to admit those persons for whom program efforts can be most helpful and screening others for referral to more appropriate sources of treatment *before* admission. Consequently, a measure reflecting the extent to which clients are transferred or referred elsewhere without successfully completing treatment in your program may provide some information useful for evaluating your program's admission policies and procedures.

In order to use the formula below, the "Reasons for Discharge" coded 2, 3, and 4 on the CODAP DR form must be combined. That is, combine Discharge Reports with any of these three code numbers entered in DR Item 6 into one group and count the number of forms in that group for the category of "transfers."

TR'_t = number of Discharge Reports submitted at the end of the month for clients transferred or referred outside the program (or clinic) (DR Item 6 codes 2 plus 3 plus 4).

T' = total number of Discharge Reports submitted at the end of the month (CFS Item 16 plus CFS Item 17).

(CODAP) Transfers as a Percentage of All Terminations

$$= \frac{TR'_t}{T'} \times 100 .$$

"All Other Terminations" as a Percentage of All Terminations

The last measure presented above provides an illustration of the potential for adapting information you already collect in one form to other uses by combining information and then following the same logic that underlies all the formulas in this manual. As you gain experience in the use of these measures and in the process of evaluation, you may want to create your own measures by separating the categories that are presented in combination in this manual or by combining them differently.

One situation that suggests such adaptation arises because several of the reasons on the Discharge Reports are similar (in that they are less related to program efforts) and because the number of terminations in a month may be so small that eight separate reasons for termination will convey a false sense of precision. As a result, you may find it more useful to disregard some of the distinctions between "Reasons for Discharge." In recognition of this possibility, we have distinguished slightly different "Reasons for Discharge" by maintaining the two most useful categories on the CODAP DR form unchanged and then combining the remaining six reasons into one broader category, according to the following scheme:

1. Successful Completions: same as DR Item 6 code 1, "Completed Treatment."

2. Left Program Voluntarily: same as DR Item 6 code 6, "Client left before completing treatment."

3. All Other Terminations: combines DR Item 6 codes:
   - 2 - Transfer to another CODAP Clinic within Program.
   - 3 - Transfer to a non-CODAP Clinic within Program.
   - 4 - Referred outside Program.
   - 5 - Program decision to discharge Client for noncompliance to Program rules.
   - 7 - Incarcerated.
   - 8 - Death.

The categories listed above are suggested because: (1) the "Successful Completions" and "Left Program Voluntarily" types of terminations should always be distinguished since they are fairly direct indicators of program effectiveness; (2) for the DARP programs, the six reasons combined here into the category "All Other Terminations" *individually* include only very small numbers of clients; and (3) the ambiguities involved in interpreting most of the six reasons separately in terms of program effectiveness decrease their value for evaluation (see the discussions above for the measures "Rule-Breakers [and] Transfers as a Percentage of All Terminations"). Based on this suggested collapsing of CODAP "Reasons for Discharge," the formula below can be used to calculate values for a measure of the proportion of terminations over which the program has less control.

S'_t = number of Discharge Reports submitted at the end of the month for clients considered to have successfully completed treatment (DR Item 6 code 1).

V'_t = number of Discharge Reports submitted at the end of the month for clients classified as having voluntarily left before completing treatment (DR Item 6 code 6).

T' = total number of Discharge Reports submitted at the end of the month (CFS Item 16 plus CFS Item 17).

(CODAP) "All Other" Terminations as a Percentage of All Terminations

$$= \frac{T' - (S'_t + V'_t)}{T'} \times 100 .$$

It should be noted that the term "All Other Terminations" is equal to the sum of Discharge Reports coded 2, 3, 4, 5, 7, and 8. Hence, the numerator in the formula above can be calculated in either of two ways:

1. As indicated, by subtracting from the total number of terminations the number of Discharge Reports coded 1 and 6 under "Reason for Discharge" (DR Item 6), or

2. By summing the numbers of Discharge Reports coded 2, 3, 4, 5, 7, and 8 under "Reason for Discharge" (DR Item 6).

Relationships Between the Termination Measures

As you may have already determined from the discussion above, the terms from several of

the termination formulas are related in the following manner:

$$T' = S'_t + V'_t + [T' - (S'_t + V'_t)] .$$

This equation says that the total number of terminations (T') is equal to the sum of the numbers of "Successful Completions" (S'_t), clients who "Left Program Voluntarily" (V'_t), and "All Other Terminations" [$T' - (S'_t + V'_t)$]. As a result, the value of any one of the measures associated with these categories of terminations can be computed by subtracting from 100 the sum of the values of the other two related termination measures.

Another relationship exists between the "Gross Termination Ratio" and all the "Reason-Specific Termination Ratios" (i.e., all the formulas with T' as the denominator). Any "Reason-Specific Termination Ratio" can be reexpressed as a percentage of all clients in treatment (rather than as a percentage of all terminations) by multiplying its value by the value (for the entire program or for the same subgroup) of the "Gross Termination Ratio." For example, the formula suggested earlier for evaluating the extent to which clients voluntarily leave the program before completing treatment is expressed in terms of a proportion of all terminations. Since, for many programs, all clients in treatment *can* leave treatment voluntarily *at any time*, it may be useful to know the proportion who do actually quit. Rather than having to compute from scratch, using a new formula, a measure of the proportion of all clients in treatment who voluntarily leave, you can determine the value of such a measure by multiplying together the values for the same time period of your program's "Gross Termination Ratio" and the measure called earlier "Left Voluntarily as a Percentage of All Terminations." The values for these two measures can be multiplied directly because the formulas for the measures (see pages 29, 31) are related in the following manner:

$$\left(\frac{T'}{N'_t} \times 100\right) \times \left(\frac{V'_t}{T'} \times 100\right) = \frac{V'}{N'_t} \times 100 .$$

Hence, multiplying any "Reason-Specific Termination Ratio" measure which has a T' as its denominator will result in expressing the "Reason for Discharge" category (or combination of categories) as a percentage of all clients in treatment *at any time* during the period.

OTHER MEASURES

Several other easily computed measures that may be useful for evaluating your program are suggested here. Not all of them are client-outcome measures, nor is the information required for calculating all of them available from the forms you currently submit for CODAP. But they are presented here primarily to challenge you to use your imagination and knowledge of your own program, information you may already have available in client records, and the simple computational techniques in this manual to devise any kind of measures that can help you learn more about your program and that can be useful in evaluating and improving the treatment of drug abuse.

Percent Living With Spouse
(Marital Stability)

Some programs are often asked by public officials, funding sources, or other interested persons about the program's effectiveness in stabilizing the lifestyle of drug abusers. While no single measure can be devised that will unambiguously provide such information, several measures can be created which can be interpreted, with qualifications, as indicators of "social stability." Probably the most important qualification on interpretations of measures of social stability is the fact that, even if drug use ceases, the lifestyles of many drug users include conventions and features that nonusers (especially "straights") would nevertheless consider socially unstable, perhaps even unacceptable.

If you believe, as many people do, that one part of the problem of drug abuse is the fact that marriage and family life is disrupted and if one of your program goals is to eliminate such familial instability, the following formula may be useful in evaluating your program's effectiveness in meeting that goal. Obtain the required information from counselors. Be cautious in interpreting results because, while drug abuse can wreck marriages, bad marriages may also drive persons to abuse drugs. Notice also that the measure is concerned only with persons *legally married*, because of the ambiguities involved in defining "consensual unions" as a stable marital arrangement.

L = number of legally married clients in treatment *at any time* during the period who each were living with their own spouse at the end of the period.

M = total number of legally married clients in treatment *at any time* during the period.

Percent Living with Spouse

$$= \frac{L}{M} \times 100 .$$

Percent Changing Place of Residence

Another possible measure of social stability is permanence of residence. While the reasons for changing are most important, a measure of only the extent of residential mobility may be helpful in judging whether drug abuse treatment needs to address different problems as well. Interpretation of such a measure is possible only after considering the reasons for changing residence. Again, the required information must be specially collected from all sources available.

G = number of clients reported to have changed residences (moved geographically) during the period.

N = total number of clients in treatment *at any time* during the period, including clients discharged or admitted during the period, for whom information about residential mobility during the period is available.

Percent Changing Place of Residence

$$= \frac{G}{N} \times 100 .$$

Percent Changing Jobs

Employment has already been suggested as an indicator of treatment effectiveness. Having a job is, of course, crucial; but staying with a job once employed may also be an indirect indicator of whether a program is continuing to have beneficial effects. Like the measure of residential mobility, interpretation of a measure of job stability *must* be based on the reasons for changing jobs; after all, a high value resulting from all the job changers moving to *better* jobs is an indication of program success as much as is a low value resulting from all employed clients staying in their present jobs. Collect the information required to compute this measure from your clients or counselors.

J = number of full-time employed clients in treatment *at any time* during the period who each changed jobs at least once during the period.

E_f = total number of clients employed full time *at any time* during the reporting period.

Percent Changing Jobs

$$= \frac{J}{E_f} \times 100 .$$

Average Time in Treatment

We have already presented some termination measures whose values for any one period are obviously meaningful and which, taken together, constitute the minimum amount of information needed to begin to evaluate your program's treatment effects in terms of termination outcomes. A different way to use information you already collect to reveal more about your program effectiveness is to summarize the termination outcomes that were expressed more specifically by those measures presented earlier. One kind of summary measure of termination disregards "Reasons for Discharge" and expresses all terminations as time-in-treatment. The degree to which the average time-in-treatment of all discharged clients approximates the term of treatment formally specified in the treatment regimen as necessary for "Successful Completion" is a rough indication of the effectiveness of your program efforts. Low values may mean that many clients are leaving treatment too early for the program to be most helpful.

The information required for the formula below can be found on your monthly CODAP Discharge Reports and Client Flow Summary form. It is suggested that you express the clients' times-in-treatment in weeks, since a month may be too long a period to reflect accurately the extent of short-term treatment while using days is unnecessarily (and tediously) precise. In order to compute the total time-in-treatment for all discharged clients, separately sum the number of months and weeks in treatment (Item 10 of the Discharge Reports submitted at the end of the month), multiply the sum of months by 4.33 (the average number of weeks in a month), and then add the result to the sum of the number of weeks in treatment. (Notice that the values for this measure are *not* expressed as percentages.)

W'_t = total number of weeks of uninterrupted treatment prior to this termination for all clients whose Discharge Reports are submitted at the end of the month (sum, in weeks, of all DR Item 10).

T' = total number of Discharge Reports submitted at the end of the month (CFS Item 16 plus CFS Item 17).

(CODAP) Average Time in Treatment

$$= \frac{W'_t}{T'}$$

This measure was not included in the section on termination measures because: (1) we consider it "optional"--a way to expand the self-evaluation process--while most of the earlier measures are "essential"--required to conduct self-evaluation; (2) its interpretation for any one month is less obvious, requiring that the other measures also be computed and useful primarily as a way of summarizing the other measures over time (see the section below on trends); and (3) it is appropriate that this measure conclude a section devoted to illustrating ways you can create your own additional measures of program effectiveness.

PART 3
COMPARISON GROUPS AND PROCEDURES FOR SELF-EVALUATION

CHAPTER VI
SUBGROUPS AND COHORTS

SUBGROUP DISTINCTIONS

It was stressed earlier (see page 11) that certain comparisons with the same measures cannot legitimately be made, and the point was illustrated by contrasting the different goals of drug abuse treatment organizations that use differing treatment modalities. The reason for distinguishing treatment modalities and computing measures for them separately is that there are important differences between modalities in treatment strategies, clientele, problems faced, goals, and objectives, among other things. Computing an overall average value for a measure forces one type of program to compare itself with a combination of all types, disregarding the differences between the modalities. If the differences are unimportant, computing the measures separately for each modality will result in very nearly equal values, so no harm is done. But if the differences between modalities *are* important, a value computed for all the modalities combined will hide that fact. Therefore, we have strongly urged that you compute values on the measures for treatment modalities separately.

The literature on drug abuse and treatment has suggested that some other distinctions besides treatment modality are also important for expectations about outcomes. Some of these important distinctions are: *intake drug use pattern, race-ethnicity, age,* and *sex,* listed here in order of their importance. It is not possible for you to change these characteristics of your clients. However, if you find in your program that certain subgroups (i.e., combinations of these characteristics) have values on the measures different from other subgroups, it *is* possible for you to tailor your program to the special needs of these subgroups. But first, it is necessary to describe a way each of the characteristics mentioned above can be used.

Treatment Modality and Environment

Following our own suggestion, we offer here one way to distinguish between treatment modalities:

1. Outpatient Methadone Maintenance (MM).
2. Therapeutic (Residential Drug-Free) Communities (TC).
3. Outpatient Drug-Free (DFOP).
4. Outpatient Detoxification (DTOP).

The symbols (MM, TC, etc.) in parentheses above are used in this manual to talk about these subgroups by treatment modality (and environment). Other distinctions are possible, and you may wish to use different categories of treatment modality (and environment) and of the other characteristics described below.* Programs with several modalities and environments may need to add categories to the list above, which includes only the most common distinctions.

Intake Drug Usage Pattern

Treatment outcomes have been shown to be

---

*The distinctions presented in this section are not the only ones which can be made; they *are* fairly useful and will be used in this manual to illustrate the principles of subgrouping. Furthermore, in the volumes of values based on DARP and CODAP programs nationwide, distinctions identical to the ones presented here for *all* the characteristics are used. To maximize the benefits of these two additional volumes of comparison values, it is suggested that you try to use the categories presented in this manual for subgroupings in your own program.

different for persons whose initial drug usage problem differs. The following distinctions can be made in intake drug usage pattern:

1. Daily use of opiates only (DA-OP).
2. Daily use of opiates and frequent use of other drugs, including alcohol (DA-OP+).
3. Less than daily use of opiates and frequent use of other drugs, including alcohol (LDA-OP+).
4. No use of opiates but problematic use of other drugs, including alcohol (NON-OP).

Again, the symbols (DA-OP, DA-OP+, etc.) in parentheses above are used in this manual (and in the volumes of measures based on DARP and CODAP) to specify one way of subgrouping by intake drug usage pattern. These distinctions indicate primarily, but not exclusively, illicit drug usage; however, nonopiate drug usage that is licit (e.g., prescription drugs) but abusive and problematic is also included.

Race-Ethnicity

In this manual, the following distinctions have been made in racial-ethnic characteristics:

1. Black (B).
2. Puerto-Rican (PR).
3. Mexican-American (MA).
4. White, not Hispanic (W).

Age Groups

Age groupings have been created for use in this manual in the following way:

1. Under 18.
2. 18-21.
3. 22-25.
4. Over 25.

Sex

Male (M) and female (F) categories are also distinguished.

CONSTRUCTING AND COMPARING SUBGROUPS

The General Method

Up to this point in the discussion we have introduced distinctions for constructing subgroups of clients that will probably be important for most programs. As will be discussed in the two separate volumes, in order to use the tables of national comparison values you will have to construct subgroups following a particular order of the distinctions presented above and using exactly the same categories.

However, you may be interested in making some comparisons that are not possible if you follow the prescribed order in the volumes based on DARP and CODAP data. In fact, you may have too few clients to be able to consider all five of the characteristics above at the same time, yet you may still want to examine distinctions within each of the characteristics separately. For example, you may want to compare male termination ratios with female termination ratios *without* regard to intake drug usage pattern, race, and age categories. (We assume that you will always make the treatment modality distinction before considering *any* other distinctions.) Consequently, you also need a general method for constructing any kind of subgroups using any kind of distinctions in any order. This is important because you will want to be able to group people in ways that will permit you to make any comparisons that *you* may feel are needed. In effect, this will give you the flexibility to compare your subgroups' values with the prescribed subgroup breakdowns for national values in the DARP and CODAP volumes *or* to construct unique subgroups to compare their values only with each other. In summary, the way you group and subgroup clients determines the comparisons you can make.

The first step in constructing any subgroup is to distinguish clients on the basis of the characteristics in which you are interested (e.g., modality, race, sex, etc.). A handy way to make these distinctions, at least at first, is to prepare a 3 × 5 index card for every client and list on the card all his/her characteristics (e.g., Tom Jones - MM / LDA-OP+ / W / 27 / M). Table 1 presents an alphabetized listing of the clients in a fictitious program with two modalities, outpatient methadone maintenance (MM) and an outpatient drug-free (DFOP) component. Several (though not all) characteristics of interest are specified in the table, as well as information about each client's use of illicit drugs during the month.

To create a subgrouping based on only one characteristic, separate all clients in your program into the categories of that characteristic. For example, table 2 illustrates this process for our fictitious program, based on modality subgroups (MM and DFOP). Notice that all other information about the

Table 1. Alphabetized list of clients in a program during 1 month with their characteristics specified

Client	Modality	Race-ethnicity	Sex	Illicit drug use during month
Aretha	MM	B	F	Yes
Bob	MM	W	M	Yes
Carlos	MM	W	M	Yes
Dan	MM	W	M	No
Eldridge	MM	B	M	No
Frank	DFOP	B	M	No
Jane	MM	W	F	No
Lela	DFOP	W	F	Yes
Maria	DFOP	B	F	Yes
Ruth	MM	B	F	Yes
Sam	DFOP	B	M	No
Tom	DFOP	W	M	No

Table 2. Clients in table 1 subgrouped on the basis of treatment modality

Client	Modality	Race-ethnicity	Sex	Illicit drug use during month
Aretha		B	F	Yes
Bob		W	M	Yes
Carlos		W	M	Yes
Dan	MM	W	M	No
Eldridge		B	M	No
Jane		W	F	No
Ruth		B	F	Yes
Frank		B	M	No
Lela		W	F	Yes
Maria	DFOP	B	F	Yes
Sam		B	M	No
Tom		W	M	No

clients is unchanged by the modality subgrouping, and simply follows each client to the appropriate subgroup.

To create subgroups based on two characteristics, categorize clients according to the first characteristic and then subdivide the first categories on the basis of each of the categories of the second characteristic. Tables 3 and 4 illustrate the procedure for this two-level subgrouping, using sex (table 3) and race-ethnicity (table 4) as the characteristics distinguished within categories of treatment modality.

You may continue this process until you have created the kinds of subgroups you want to make, *or* you may use the same characteristics but change their ordering to create

Table 3. Clients in table 1 subgrouped first on the basis of treatment modality, then on the basis of sex

Client	Modality	Race-ethnicity	Sex	Illicit drug use during month
Aretha		B		Yes
Jane		W	F	No
Ruth		B		Yes
	MM			
Bob		W		Yes
Carlos		W	M	Yes
Dan		W		No
Eldridge		B		No
Frank		B		No
Sam		B	M	No
Tom		W		No
	DFOP			
Lela		W	F	Yes
Maria		B		Yes

Table 4. Clients in table 1 subgrouped first on the basis of treatment modality, then on the basis of race-ethnicity

Client	Modality	Race-ethnicity	Sex	Illicit drug use during month
Aretha			F	Yes
Eldridge		B	M	No
Ruth			F	Yes
	MM			
Bob			M	Yes
Carlos		W	M	Yes
Dan			M	No
Jane			F	No
Frank			M	No
Maria		B	F	Yes
Sam			M	No
	DFOP			
Lela		W	F	Yes
Tom			M	No

different subgroups. As mentioned earlier, however, we recommend that you always differentiate between treatment modalities *first*, regardless of the order in which you use the other characteristics.

Minimum Size of Subgroups

You may have noticed that, as the subgroup detail increases, the number of persons in a subgroup decreases, sometimes to just one or two, occasionally to zero. As a rule of thumb, *no value for a subgroup which contains fewer than five persons can be relied upon.* Within your own program, this rule of thumb also applies. Therefore, if you make

distinctions in such detail that one or more of your subgroups dwindles below five members, you've gone too far. Two techniques can be used to avoid this problem.

Collapsing Categories

The first technique you should try is the use of more crude categories within characteristics. Thus, instead of using the four age groups we suggested, combine them and use only two (e.g., Under 25, Over 25). Even so, you may still have subgroups that are too small, in which case that characteristic has to be ignored. (Even though you may not be able to make distinctions concerning a characteristic, that factor still exists and may be important--the problem is that you cannot examine its importance.)

Limiting the Number of Levels

The second technique involves making subgroups by considering only one or two characteristics at a time. Using this technique still allows you to make many important comparisons, although other interesting and possibly important comparisons cannot be made using this technique. Nevertheless, the reliability of the values you can compute for your own subgroups is more important than the detail of the breakdowns. And even though such limited subgroups do not allow you to control for all factors at once, you can still reliably compare with each other the subgroups you *can* construct to determine whether your program is serving them all equally well.

Computing the Measures for Subgroups

Once you have separated all your clients into subgroups based on one or more characteristics, record for each client the information needed to compute the measure on which the subgroups will be compared. Subgroups such as Housewife (H), Student (S), and Employed Full-Time (Ef) would be necessary for computing the measure "Percent Employed Full-Time" (DARP). Then, within each subgroup, count and record the total number of clients and the number of clients required by the formula categories for the measure. After this has been completed, compute the value for *each* subgroup, using the formula for the measure in which you are interested.

Table 5 summarizes the information needed to compute the measure "Percent Drug-Free (Outpatient)," using the DARP formula from page 25, for all the example subgroups in tables 1 through 4 above.* Computation of the values in table 5 is fairly straightforward, resulting from counting the appropriate clients in the subgroups in tables 1 through 4, then substituting the numbers in the DARP formula for the measure "Percent Drug-Free. You may find it a useful exercise to go back to tables 1 through 4 and trace the process that resulted in table 5.

Table 5. Information for computing "percent drug-free (outpatient)" and value for each subgroup in tables 1–4

Table No.	Subgroups	D_i	N_d	$\left(\frac{N_d - D_i}{N_d} \times 100\right)^*$
1	None	6	12	50%
2	MM	4	7	43%
	DFOP	2	5	60%
3	MM/F	2	3	33%
	MM/M	2	4	50%
	DFOP/M	0	3	0%
	DFOP/F	2	2	100%
4	MM/B	2	3	33%
	MM/W	2	4	50%
	DFOP/B	1	3	67%
	DFOP/W	1	2	50%

*This formula is the DARP formula for the measure "percent drug-free (outpatient)" discussed on page 25.

Notice in table 5 that the sizes of subgroups (i.e., the numbers in the column labeled N_d) from tables 3 and 4 are smaller than our rule of thumb number of 5. While the computations of the measure for these too small subgroups in table 5 are, therefore, technically unreliable, we have nevertheless performed the calculations for two reasons: (1) to continue the illustration of constructing subgroups and calculating their values on the measures and (2) to illustrate the reason for our caution about subgroups with fewer than five clients. Consider that, for a subgroup of 5 (e.g., DFOP in table 5), a change of one client in the D_i column will result in a 20 percent

*We present the information and compute the measure for the *entire* fictitious program solely to illustrate the mechanics of computing values for subgroups. Consistent with our earlier suggestion, values that disregard differences in treatment modalities should *not* be computed for actual programs.

increase or decrease in that subgroup's "Percent Drug- Free" value. As the size of a subgroup decreases to less than 5, changes of one client in the D_i column result in changes of 25, 33, 50, and 100 percent in the value of the measure for each subgroup that becomes successively smaller by one. Changes of such magnitude in the value of a measure provide indications that are increasingly out of proportion to the actual treatment effort (or lack of it) which resulted in changes of only single clients. Expressed differently, evaluation results will be affected (favorably *and* unfavorably) more by the actions of a single client in small subgroups than in larger subgroups, even though the effect (or neglect) involved in changing a single client may be very nearly the same.

A final detail must be introduced concerning the dependence of the computation of the measures on the way the subgroups are constructed. With the introduction of each additional characteristic, the measures for the more detailed subgroups *must* be recomputed, because at each successive level of detail the new arrangement of characteristics requires that clients be reassigned to different subgroups. This reassignment almost always results in a rearrangement of clients within the formula categories for the measures. For example, it can be seen in table 5 that there is no relation between the "Percent Drug-Free" values for the modality/sex subgroups and the values for either the modality-only subgroups or the entire program.

Less apparent from table 5 is the fact that we often construct subgroups based on two or more characteristics (e.g., MM/F and MM/M) without pausing to compute the measures at the intermediate levels (e.g., MM). For the same reason that measures have to be computed anew for each additional subgrouping characteristic, they must also be recomputed when subgroups are combined. Often, detailed subgroups are combined by eliminating a characteristic and the values for the eliminated subgroups are averaged on the *mistaken* assumption that the average of the lower-level values will equal the intermediate-level value. For example, eliminating the sex characteristic from the four subgroups (MM/F - MM/M and DFOP/M - DFOP/F) in table 5 will reconstruct the modality-only subgroups, MM and DFOP. But averaging the "Percent Drug-Free" values for the two sets of same-modality/different-sex subgroups will *not* result in the correct values for the MM and DFOP subgroups.

MM/F	=	33%	
MM/M	=	50%	
		83%	÷ 2 = 41% (average MM)
DFOP/M	=	0%	
DFOP/F	=	100%	
		100%	÷ 2 = 50% (average DFOP)

Neither value is correct, although the "average MM" value is closer to being right (correct MM = 43 percent) than is the "average DFOP" value (correct DFOP = 60 percent). The rule is clear: the value on a measure for any subgroup at any level must be computed from the assignment of clients in that subgroup; no subgroup value can be computed by dividing any higher-level value or by combining any lower-level values.

Comparing Subgroup Values on the Measures

Based on only the five characteristics (modality, intake drug use pattern, race-ethnicity, age, and sex) and the distinctions within them that we have discussed above, thousands of subgroups can be constructed. Even within a single modality, hundreds of subgroups are possible, simply by changing the order of the characteristics.* Nobody, least of all program staff and not even researchers, has the time or patience to construct, compute measures for, and compare all these possible subgroups. Fortunately, most of the possible subgroupings are also trivial (i.e., unimportant because repetitious or unnecessary). The trick is to identify the important and, therefore, useful, subgroupings. Three techniques are helpful for this purpose.

*It is for this reason that "researchers" use *theory*. Theory provides a logical, consistent, and systematic explanation of events. Based on such an explanation, a researcher is led to expect certain characteristics to be important in certain ways, while different orderings or characteristics are expected to have no effect. By paying attention only to the characteristics and their ordering suggested by a theory, the researcher saves time and effort; if the data analysis produces expected results, he comes to have more faith in his theory; only if the predicted effects are *not* demonstrated does the researcher reorder his characteristics or introduce new factors.

Use Knowledge From Other Sources

When we first introduced the five characteristics for distinguishing subgroups, we cited the fact that research has shown them to be among the *most* important, *in the order presented*, for identifying subgroups whose drug abuse problems seem to require special treatment attention. Therefore, we recommend that you tentatively plan to use these distinctions, following the suggested ordering as far as possible, to determine whether research findings elsewhere also apply in your program.

Investigate Each Characteristic Separately

The first step you should undertake in determining which subgroups will be most useful for self-evaluation is to construct the second-level (after modality) set of subgroups based on only one characteristic at a time.* With few exceptions, if values on the measures do not differ for subgroups based on a single characteristic, that characteristic will also not be important in combination with other characteristics, regardless of ordering. For example, in table 5, the subgroups MM/W and DFOP/W both have values of 50 percent on the measure "Percent Drug Free." In this example, the only difference between the two groups is in treatment modality. Consequently, the equal values indicate that different treatment modalities are not related to the use of drugs by whites. It is possible, therefore, to disregard modality for the white clients in tables 4 and 5, recount the D_i (=3) and N_d (=6) categories, and recompute a value of 50 percent again for the subgroup W. Of course, because the drug abuse behavior of blacks *does* differ depending on treatment modality and because there *is* a difference between whites and blacks in the same modalities, eliminating the MM and DFOP distinction for whites serves only to help you meet the minimum subgroup size requirement. But this example illustrates how you can economize on subgroup detail:

1. Compute several measures for all the subgroups created with only one characteristic; and
2. If there is no difference for each measure between the values for all the subgroups, you can disregard that characteristic.

This section has been concerned with the value of making subgroup breakdowns. The discussion so far has proceeded on the assumption that values on outcome measures vary by modality, intake drug usage pattern, race-ethnicity, age, and sex groupings. In your program, however, one or more of these distinctions may result in subgroup values which are essentially equal. *Such findings mean only one thing: you do not need to bother with any subgroup distinctions resulting in equal values.* While the *necessity* is thus eliminated for considering these "inconsequential" characteristics any longer, you may want to retain them in order to allow comparison with values from the national DARP and CODAP files. If you do discard these "inconsequential" distinctions, some of the usefulness of the volumes containing values from nationwide DARP and CODAP programs will be lost, especially if the eliminated factors are race-ethnicity or intake drug usage pattern. But this disadvantage is compensated (often, *more* than compensated) by the increase in the reliability of values for the less-detailed subgroups you continue to distinguish, due to the increase in their membership size and by the savings in time and effort formerly devoted to constructing and computing values for the unneeded subgroups.

Appropriate Versus Possible Comparisons

Neither construction nor comparison is appropriate for all the hundreds of possible subgroupings. The two preceding sections have been devoted to discussing ways to limit the *construction* of subgroups only to those that are useful. This section focuses on limiting the number of *comparisons* that need to be made, even after limiting the number of subgroups constructed. Some comparisons are obviously inappropriate. For example, you may find after investigating each characteristic separately that values of your outcome measures vary by modality, intake drug use pattern, and sex, resulting in the need to construct all the possible subgroups based on these three characteristics.

Two subgroups that can be distinguished at the two-characteristic level are MM/NON-OP and MM/F, both based on a common modality distinction, but differing on the characteristic used to make distinctions at the second level. While it is possible to

*Again, we assume that you will automatically differentiate between treatment modalities, so the strategy presented here applies to the other four characteristics we suggested and any others in which you are interested.

compare values on a measure for these two subgroups, it makes no sense to do so, as it cannot be determined whether differences in subgroups values are due to distinctions within the second-level characteristics or to the fact that the second-level characteristics are themselves different. In other words, "you shouldn't compare apples and oranges," even though you can.

Of the remaining comparisons that are appropriate, you should be selective so as to make only as many subgroup comparisons as will be most useful for evaluation. The general principle to follow is to compare only subgroups with identical characteristics broken down in the same order. Again referring to table 5, there are many appropriate comparisons. For example, comparing "Percent Drug-Free" values for the subgroups MM and DFOP seems to indicate that the drug-free modality is more effective than methadone maintenance, as expected, although treatment outcomes for clients in the drug-free modality are considerably below the goal of 100 percent drug-free. Further, comparisons of the subgroups MM/F with DFOP/F and MM/M with DFOP/M help us specify the original difference between modalities: not only do *all* clients in different modalities experience different outcomes but the females seem to do better in a drug-free modality, while males do better in methadone maintenance. On the other hand, comparisons of the subgroups MM/F with MM/M and DFOP/M with DFOP/F provide indications of the differential outcomes experienced by males and females in the *same* modalities (i.e., while holding the effects of modalities constant). For the same subgroups, other comparisons are appropriate (e.g., MM/F and MM/M with MM), but unnecessary because the same information just discussed would be revealed. The subgroups based on modality and race-ethnicity can also be compared among themselves and with the modality-only subgroups in a similarly selective manner, so that all the information is revealed without having to make every possible comparison.* In this way, you can maximize the value of subgroupings while minimizing the time and effort required to use them in evaluation.

*You may find it useful to practice making comparisons between the modality/race-ethnicity subgroups and their "Percent Drug-Free" values in table 5, checking your procedure against our discussion here of the modality/sex subgroup comparisons. It should be noted that comparisons of any modality/sex subgroups with any modality/race-ethnicity subgroups are clearly *not* appropriate, since the comparison of sex with race-ethnicity cannot be interpreted.

COHORTS: TIME-BASED GROUPS

Introduction

The measures and the new method of grouping clients to be presented in the following sections are slightly different from those you have already seen in this manual. The difference is really only in the way you arrange the items of information about clients *before* you begin your computations, not in the difficulty of the computations themselves. As a matter of fact, three of the measures for which cohort values are provided in the volume based on DARP programs use exactly the same formulas presented earlier for the computation of program and subgroup values.

Because information has to be rearranged and because the same measure has to be computed for more groups, using cohort methods will require some additional time and work. But the extra time and work also result in extra benefits: *you gain more detailed information about your program's status and history and you can begin to assess the outcomes of treatment at various stages during a treatment career in your program.*

At this point, it is necessary to define a cohort, to tell you how to create one kind of cohort, and to discuss the value of using such time-based groups.

Definition of a Cohort

In general, a cohort is "an aggregate of individuals who experienced the same event (such as admission to treatment) within the same time interval" (see page 14). For our purposes, a more narrow definition is most useful: a cohort is a group of individuals who have been in a treatment program *for the same length of time*. Thus, a 1-month cohort is all clients who have been in treatment for any part of 1 month but no more, while a 2-month cohort is all clients who have been in treatment for at least 1 month but no more than 2 months.

It should be noted that an individual can only be included in one cohort; i.e., although a person in the 2-month cohort has also been in treatment for 1 month, he is not also included in the 1-month cohort. Rather, we use cohorts and compute them just like we use and compute a person's age: a 25-year-old person has also been all ages

below 25, but he always has only one age--his latest. As a matter of fact, one way we use and compute our ages is an example of a specific type of cohort--groups of individuals who were born within the same calendar year; i.e., "birth" cohorts.

How to Create Cohorts

While there are many events on which to base cohorts, we will focus here on only one method of constructing cohorts. This method was chosen because it provides the most useful information for treatment evaluation for the effort involved. The method we recommend and have used in the remainder of this manual defines all clients admitted to treatment during each single (calendar) month as a separate cohort. That is, all the clients entering treatment *at any time* during the same calendar month *are* the cohort for that calendar month. Clients continuing in treatment for more than 1 month remain assigned to the cohort for their admission month. Thus, the "June 1976 Cohort" is constructed of all clients admitted at any time during the calendar month of June 1976 and is restricted thereafter to that original set of clients.

Cohorts Compared With Subgroups

At first glance, there may seem to be little difference between subgroups and cohorts, and, indeed, the method of constructing both is similar. After identifying the characteristic of interest (e.g., sex for subgrouping, month of admission for cohorts), you separate the clients into categories of the characteristic (e.g., sex--M and F; cohorts--month of admission). To this point, there is no difference between the two methods.

Subgrouping *requires* that clients admitted each month be added to subgroups already formed (e.g., newly admitted males must be *added to* the group of males already in treatment) or that a client whose status changes be transferred to a different subgroup (e.g., after his 22nd birthday, a client must be *reassigned* from the "18-21" age group to the "22-25" age group).

But because cohorts are "time-bound," once a cohort is constructed, no new members can be added, nor can any of the original members be reassigned to another cohort as long as they are continuously in treatment. For example, beginning July 1, 1976, no new clients can ever again be admitted to the "June 1976 Cohort" because the calendar month June 1976 will never occur again, nor can a client admitted during June 1976 drop back to the "March 1976 Cohort" or advance to the "August 1976 Cohort." Readmissions represent a complication of this system. We recommend that they be treated like new clients upon readmission, as though they had never had any treatment (though for a client with a long time-in-treatment who drops out for a short period, you may be justified if you reassign him to his original first-admission cohort). The only change in membership that *can* occur for cohorts is loss of terminated clients (discussed below). Admission to treatment is analogous to birth--no one can be born in any year other than their actual year of birth.

Time-in-Treatment

Once constructed, cohorts can be *numbered* according to the whole number of months that includes all the time each cohort's clients have been in treatment, as of the first day of the current month. Length of time-in-treatment for every client in the same cohort is determined by subtracting the cohort's admission month from the first day of the current month and year. For example:

March 1976 cohort:

Current month and year	= May '76	= 5th month, 76th year
Admission month and year	= March '76	= 3rd month, 76th year
	Difference =	2 months, 0 years

In this example, clients admitted during March 1976 belong to the "March 1976 Cohort," which can be numbered Cohort #2 on May 1, 1976, to signify that the cohort members have been in treatment for 2 months. On June 1, 1976, the same cohort becomes the 3-month cohort or Cohort #3. With each successive month, every cohort becomes one month-in-treatment older and the numbering increases by one for every cohort.

Observe that, in the example for the "March 1976 Cohort" above, the current year and the admission year are identical. If the current year differs from the admission year, the current year must be converted to be equal to the earlier admission year by transferring 12 months for every year of difference to the "month" column. For

example:

November 1975 cohort:

Current month and year	= May '76 = 5 + 12	=	17th month, 75th yr.
Admission month and year	= Nov. '75	=	11th month, 75th yr.
	Difference	=	6 months, 0 years

Notice that this results in the cohort numbering discussed in the previous paragraph. Such a conversion is necessary to express the answer in months, the only numbering system used for the cohorts. We have no cohort periods expressed in years, even though there may be clients in treatment for 12 months or longer; such cohorts are expressed as the "12-month cohort" (Cohort #12), the "13-month cohort" (Cohort #13), etc.

In order to maintain the integrity of the "month" time period (that is, in order to assure that all our cohorts cover an equal and approximately 30-day period), cohorts' time-in-treatment must be computed (or renumbered, as discussed below) *as of the first day of the month every month.* Thus, you can construct cohorts and compute their time-in-treatment only for *full calendar* months that have elapsed; i.e., only for months for which you have the full calendar month's data. Since you have not completed all of the current month, you cannot construct a cohort for any part of it, so long as your cohorts are based on calendar *months*. Thus, when you work with cohorts, you must disregard any clients admitted or discharged during the current month *and any information gathered during the current month on your continuing clients*; however, all this current information will be used for the next month's computation. For example, measures for the *cohort* of clients admitted during March cannot be computed until after March 31 because the *cohort* cannot be constructed until its "month" is over.*

If you have gotten this far, you may have imagined how much work is involved in assigning clients to cohorts every month and you may be discouraged when you anticipate having to repeat the operation every month. TAKE HEART! You do not have to go through the entire procedure of assigning each client already in treatment to a cohort every month. Admittedly, for the first month you must assign clients to their admission month cohorts and calculate the difference between each cohort's admission month and the current month. But after that first month, you simply increase each cohort number by one (+1) for each successive month any clients in that cohort remain in treatment, including the month in which the last cohort member is terminated.

Removing Terminated Clients From Cohorts

We have seen that *all* admissions during a calendar month are assigned to Cohort #1 as of the next month's first day, regardless of the actual number of days in treatment during the admission month. On the other hand, a termination is *recorded* for a cohort *in* the calendar month in which it occurs (i.e., *as of the first day of the next month* after the month of termination). *But*, the terminated client's name and record is *not eliminated until the first day of the second calendar month* following the month of termination. You *must* maintain your cohort records in this fashion so that you will be able to credit terminations to the proper month and include clients in the computation of measures for every month treatment was received, yet still be able to eliminate terminated clients before they get included in calendar month time periods during *no* part of which they were in treatment.

Table 6 provides an example of all these instructions for constructing and maintaining cohorts in a fictitious program that started assigning clients to cohorts on April 1. Thus, once you first assign clients

Table 6. Constructing, maintaining, and dropping clients from cohorts in a program

Client	Date admitted	Date discharged	Months-in-treatment (Cohort No.) as of: 3/1	4/1	5/1	6/1	7/1
Bob	Jan. 5	–	–	3	4	5	6
Jane	Jan. 31	Apr. 1	–	3	4	–	–
Daniel	Mar. 30	–	–	1	2	3	4
Maria	Apr. 15	–	–	–	1	2	3

*Information about clients admitted, discharged, or continuing treatment during a month *can* be used to compute that same month's measures for the program or for subgroups constructed as discussed previously in this chapter.

to cohorts, all you have to do *as of the first of each month* is renumber the old cohorts, assign clients admitted the previous (*not* the current) month to a *new* cohort (#1), record terminations during the previous month, and eliminate records of clients who terminated the month *before* the previous (not *during* the previous) month. This method is similar to what we all do each year to calculate our ages--we add 1 year to our previous age on our birthday, as a shortcut to recomputing the difference between the current date and our birthdate; the major differences are that we do not tie our ages exactly to calendar years and we don't usually go back after a death to compute measures.

Cohort Worksheets

Now that we've discussed how to create cohorts and shown that, once created, cohorts are fairly simple to maintain, a suggestion is offered here to try to help minimize the paperwork, time, and effort involved in keeping up with cohort records. While there is no way at the outset to avoid reviewing each client's clinical records and calculating each client's cohort assignment, use of a worksheet similar to the one illustrated below can make collecting the necessary information more systematic and simplify the task of maintaining such records. (A copy of two suggested worksheets, complete with accommodations for all the information needed for the formulas discussed in this manual, are included in appendixes A and B.) Notice that we have included in the illustrated worksheet space for recording treatment modality, intake drug usage pattern, race-ethnicity, age, and sex for each client so that reference does not have to be made each month to the clinical records for this information.* Figure 6 provides samples of partial worksheets for maintaining records for two cohorts.

To begin, observe that every cohort is identified in two ways: (1) by admission month and year (e.g., the "January 1976 Cohort"), and (2) by a number denoting the time-in-treatment of all clients belonging to the cohort (e.g., "Cohort #5" or the "5-month cohort"). For the remainder of our discussion, we will normally use the "Cohort #5" or the "5-month cohort" alternative, rather than referring to cohorts by their admission month and year. Also, after you assign clients to cohorts, you no longer need to know during what month they were assigned, so there is no provision on the

*However, we recommend that you periodically check the clinical records or that you enlist the aid of your counselors to assure that changes in treatment modality and age are recorded.

WORKSHEET FOR CODAP FORMULAS — January '76 COHORT # 5 (MONTHS IN TREATMENT AS OF June 1976 CURRENT MONTH)

CLIENT NAME	ADMISSION DATE	DISCHARGE DATE	TREATMENT MODALITY: MN	TC	DFOP	DTOP	INTAKE DRUG USE PATTERN: DA-OP	DA-OP*	LDA-OP*	NON-OP	RACE-ETHNICITY: B	PR	MA	W	AGE GROUP: UNDER 18	18-21	22-25	OVER 25	SEX: M	F	EMPLOYMENT: E1	E2	U	O
Bob	1-5-76		X				X							X		X			X			X		X
~~Ruth~~	~~1-28-76~~	~~4-15-76~~	X																					
Eldridge	1-29-76		X																					
~~Jane~~	~~1-31-76~~	~~4-1-76~~	X																					

WORKSHEET FOR CODAP FORMULAS — March '76 COHORT # 3 (MONTHS

CLIENT NAME	ADMISSION DATE	DISCHARGE DATE	TREATMENT MODALITY: MN	TC	DFOP	DTOP	INTAKE DRUG USE PATTERN: DA-OP	DA-OP*	LDA-OP*	NON-OP	RACE-ETHNIC: B	PR	MA
Sharon	3-1-76		X				X						
~~Carlos~~	~~3-10-76~~	~~4-5-76~~	X					X				X	
Aretha	3-27-76	5-28-76	X						X		X		
Daniel	3-30-76		X						X				

Figure 6. Sample worksheets for maintaining cohort records (Current month and year is June 1976 for both.)

form for such information. But it is a good idea to keep a record of the "Current Month" and "Current Year" as you change the "Cohort #," in case you cannot get to your computations one month; upon returning to your worksheets, you will know immediately that the "Cohort #" is incorrect and how to update the worksheets.

Now, to minimize the work in maintaining cohort records, all that is necessary, using these worksheets, is each month to change the number in the title (following "Cohort #") and change the "Current Month" and "Current Year" designations. Thus, for our example worksheets, once the "Current Month" becomes July, we need only erase "June," write in "July" on both sheets, and change the "Cohort #" to "6" on the back sheet and to "4" on the front sheet. You must also eliminate clients in the cohort who terminated treatment during the month before the last month. Thus, in our example worksheets, Ruth, Jane, and Carlos all terminated during April and their records are eliminated as of June 1. Notice that terminated clients must have their *entire* record eliminated, which we have illustrated for Carlos by drawing a line all the way across the worksheet. We recommend, however, that you erase completely all entries after the "DISCHARGE DATE" for clients dropped from cohorts to avoid any possibility of incorrectly constructing subgroups or computing measures. On the other hand, any records of clients terminating *during* May must be retained in the cohorts until July 1 although the terminations occurring in May must be recorded (e.g., see the worksheet record for Aretha).

One further item about cohorts should be noted before we leave the worksheets. The only change in membership of cohorts that *can* occur is loss of terminated clients.

The Value of Cohorts

The last topic discussed above holds the key to understanding both the nature and the value of cohorts. If you grasp the mechanics of creating cohorts, the fact that, once created, the cohorts themselves never change even though their time-in-treatment increases, and the implications of the constancy of their membership, you will be well on your way to being able to conduct and interpret sophisticated program self-evaluation.

The value of cohorts lies, first, in the fact that by comparing several cohorts at one point in time, you can make two valuable interpretations. The first of these is identical to the interpretation you make when you construct subgroups and compare their rates. That is, time-in-treatment is simply another distinction you can make, and the cohorts are the subgroups. You can thereby determine whether at any point in time, time itself (as time-in-treatment) makes any difference. Related to this conception of cohorts as subgroupings is "time-series" analysis (discussed below), in which you can identify a "profile" of your program.

The second way in which cohorts are valuable is related to the constancy of their membership. By following single cohorts through time (true "cohort analysis"), we are better able to specify the effects of treatment, not on a single individual since you already have that assessment in your clinical records, but on groups of the *same* individuals. Such an approach is equivalent to describing a "typical career" (though *not* a typical *individual*) through the program. Thus, the different responsiveness of individuals to treatment is largely cancelled, allowing the effects of program treatment to become more apparent. Cohort analysis in this more limited sense will be discussed in chapter VII.

COMPUTING MEASURES FOR COHORTS

Using Chapter V Measures
With Cohorts

All the measures previously presented can be used with cohorts and, indeed, their value increases if you compute them for both the program and subgroups and, then, for cohorts. As a matter of fact, three of the measures for cohorts in the volumes using DARP and CODAP data are measures first introduced in chapter V. To demonstrate that the earlier measures are applicable to cohorts, the computational techniques for the measure "Percent Unemployed" (DARP) (see pages 19-20) for use with cohorts is discussed here.

To compute the "Percent Unemployed" in a cohort, you must first place all clients *in that cohort* into two groups:

U_{Cn} = number of clients in "Cohort #n" who were *unemployed during the entire reporting period.* (See pages 18-21 for instructions about housewives and students.)

N_{Cn} = total number of "Cohort #n" clients in treatment *at any time* during the reporting period, including clients discharged during the period. (Note: clients *admitted* during the period are included only if the cohort is Cohort #1.)

(DARP) Percent Unemployed in "Cohort #n"

$$= \frac{U_{Cn}}{N_{Cn}} \times 100 .$$

The subscript "Cn" is used to indicate that the base unit is a cohort (C), rather than the entire program, and to identify the cohort number (n), since the measure has to be computed for each cohort separately.

The same adjustment to include only cohort members must be made to the base unit for computing *any* measure for a cohort; but after you adjust the unit on which the computation is being made, all the formulas presented in chapter V will provide values for cohorts as well.

Once you select a few of the measures introduced earlier for use with cohorts, adjust the base units to which the formulas apply, and compute the values, you must consider the limitations of the measures and what they mean for evaluating your program. As a first approximation, the limitations discussed in chapter V on each measure apply even when the measures are used with cohorts, since the measures still portray only a "cross-sectional snapshot" of a cohort. But using cohorts as the base units for taking cross-sectional views, rather than using the program or subgroups as the base, alleviates some of the *seriousness* of the limitations associated with each measure.

First, using cohorts makes it easier to eliminate some of the possible interpretations that could legitimately be attached to an overall program value for a measure, in a manner similar to the way comparing subgroups helps focus on the correct interpretation. For example, using subgroups as the base, a high program unemployment ratio is "explained" if, upon comparing racial-ethnic subgroups, you find that blacks have an unemployment ratio much higher than whites, and blacks also constitute the majority of your program's clientele. In this case, the possible meaning that the high program unemployment ratio is a consequence of a high community unemployment rate can probably be eliminated since whites can find employment. Thus, you can focus your program efforts on improving job skills of blacks or helping blacks combat racial discrimination in hiring. In a similar manner, with cohorts as the base units for the measures, you can also eliminate some of the possible meanings associated with values for a measure. But the worth of using cohorts for this purpose lies in the fact that the remaining interpretations of the measure are more likely to be program connected since cohorts are defined in terms of time in the program.

Hence, for example, a high program unemployment ratio may be found after comparison of cohort unemployment ratios to be a consequence of a disproportionately high number of new and unemployed clients who can be expected to find work as a result of staying in the program longer, *or* it might also be a consequence of unemployed persons staying in the program longer while clients successful in finding a job tend to leave the program earlier. The community unemployment situation need not even be considered to interpret the program unemployment ratio. In the first case, you would want to focus efforts on keeping the clients in the program as long as possible, being confident that they will find employment in conjunction with increased time-in-treatment, while in the second case, you would be advised to devote more effort to actually getting clients employed. These examples illustrate, then, that use of *any* measures with cohorts, while still being subject to some of the same limitations applied at the program or subgroup level, provides more program related information by helping eliminate some nonprogram explanations. *This ability to eliminate some nonprogram information with cohorts is due to the fact that cohorts are tied to the length of time a group is in the program.*

Second, computing the measures presented earlier for cohorts modifies the limitation arising from the measures being computed on different people at different times. A cohort is, with the exception of terminations, the same people measured at a later time as those who were measured at an earlier time. Any differences between successive values on any measure for a cohort, then, can be attributed to the effects of either program related factors or nonprogram factors, but not to the fact that the cohort has different individuals later than it had earlier. This relative constancy of the membership of a cohort, plus the ability discussed above to eliminate some nonprogram factors, results in measures whose meaning for your program is somewhat more obvious and less complicated by nonprogram factors.

Some Qualifications on Using Cohorts

Using cohorts as base units for part of your evaluation effort is strongly recommended, both as a means to illuminate values computed for the program and for subgroups and for the benefits to be derived directly. However, some caution must be exercised, and the benefits are gained at some cost. We've already mentioned the cost in terms of additional time, effort, and recordkeeping involved. Here, we will discuss some qualifications on the use of cohorts, some problems you may encounter, and suggestions for resolving them.

Subgroups and Cohorts

Unless you have a very large program (e.g., 500+ clients), you should not attempt to construct subgroups, then make cohort distinctions within each subgroup. The procedure is tedious, errors are easily made, and similar results can be obtained by reversing the procedure; that is, constructing cohorts and then making subgroup distinctions within each cohort. This method is especially recommended in anticipation of the finding that time-in-treatment (i.e., cohorts) will be the most important factor you are likely to consider.

Minimum Size of Cohorts and Their Subgroups

Earlier, we strongly suggested that you not even compute rates for subgroups with fewer than five clients. Because of the value of cohorts, it is permissible to bend this rule slightly for cohorts and for subgroups *within* cohorts: cohorts and subgroups within cohorts should have no fewer than *three* persons to justify computing a value for the cohort or subgroup. Accepting this reduction in size is allowable because assigning clients to time-in-treatment cohorts before subgrouping has the effect of stabilizing the resulting subgroup membership better than any other characteristic of clients. Thus, within cohorts, make no subgroup distinctions where one or more subgroups will have fewer than three members. In small programs, this restriction may prevent any subgroups from being formed when using cohorts; if so, the usefulness of cohorts exceeds the loss of subgroup distinctions within cohorts, as long as you are also computing the measures separately for the subgroups alone.

A further relaxation of the minimum-size rule applies when you are computing termination measures on cohorts. With termination measures *only*, we are as interested in the last termination as we are in the first one, since the last one sets the maximum limit on the range of time-in-treatment for the membership of any one admission cohort (which is what any time-in-treatment cohort also is). Thus, you not only can, but you should, compute termination measures for every cohort that still has any members, no matter how few.

Length of Time Periods Covered by Cohorts

In some of the smaller programs, especially those whose "prescribed" treatment term is fairly long (e.g., 6 months or more), not only will it be impossible to distinguish subgroups within cohorts if the minimum-size rule is followed, but the cohorts themselves may not have sufficient members to allow computing *reliable* rates. In such situations, you can still retain some of the advantages of using cohorts by combining your cohorts to cover 2, 3, or more months. However, unless your program requires an exceptionally long time to complete (e.g., 18 months or more), cohorts combined to cover more than 3 months may not be too useful. If you must use longer-term cohorts, you should still try to have each current admission cohort (#1) cover no more than 1 month because the first month in treatment is such a crucial time.

Using 3-month periods for illustration, one method for combining cohorts still uses the calendar month as the reporting period, and measures are computed separately for the 1-month cohort (#1) each month and for the remaining cohorts by the method of "moving averages."

First, *all* clients are assigned to the cohort for their *calendar* month of admission. There will be only one current admission cohort (#1) each month on which you compute measures, and different clients will be in each month's current admission cohort. With regard to the remaining cohorts, each one will always have the same members (except for terminations) just like all other cohorts discussed throughout this chapter. But each month, combine the current 1-, 2-, and 3-month cohorts into a single "up to 3 months in treatment" cohort, the 2-, 3-, and 4-month cohorts into a separate "2 to 4 months in treatment" cohort, and so on until all clients are in 3-month (in treatment) cohorts. Then compute measures for these 3-month cohorts using data collected (events occurring) only during the single previous month but for clients in all three of the 1-month

cohorts now assigned to each 3-month cohort. Table 7 contrasts the method discussed earlier in this chapter ("normal" method) and the present "moving average" method.

As you can see from table 7, computing measures for the current month for all cohorts involves data collected (events occurring) only during the single previous month. The "moving average" method also requires data collected only for the previous month but grouped for all clients admitted during consecutive 3-month periods. This results in overlapping two single month admission cohorts from the previous computation for every successive computation. Using the "moving average" method involves slightly more preparation before computation than the normal method, but it provides monthly information for almost as many 3-month cohorts as there are 1-month cohorts, it smooths some of the fluctuations that would occur with small-sized groups, and the overlap compensates somewhat for having to define cohorts in terms of periods longer than 1 month.

Table 7. Comparison of "normal" and "moving average" methods of constructing cohorts to meet minimum size requirements

Current month begins	Use events occurring to all clients (data) during	Method	Time in treatment	Compute measures on cohorts composed of clients admitted during					
			5	11/75					
		NORMAL	4		12/75				
		(1-month)	3			1/76			
		cohorts	2				2/76		
			1					3/76	
Apr. 1, '76	Mar. '76								
		MOVING AVERAGE	3-5	11/75	12/75	1/76			
		(3-month)	2-4		12/75	1/76	2/76		
		cohorts	1-3			1/76	2/76	3/76	
			1					3/76	
			6	11/75					
		NORMAL	5		12/75				
		(1-month)	4			1/76			
		cohorts	3				2/76		
			2					3/76	
			1						4/76
May 1, '76	Apr. '76								
			4-6	11/75	12/75	1/76			
		MOVING AVERAGE	3-5		12/75	1/76	2/76		
		(3-month)	2-4			1/76	2/76	3/76	
		cohorts	1-3				2/76	3/76	4/76
			1						4/76

CHAPTER VII
TRENDS

INTRODUCTION

In part 1 of this manual, we stressed the importance of continuing your evaluation efforts beyond one point in time, of looking at "changes" in your program values, of interpreting values "over time," and of observing any differences in values on a measure between one time and the next. Such an emphasis on time has not been intended to take anything away from the importance of values computed, either on the program and subgroups or based on cohorts, for a single point in time (e.g., a month). As a matter of fact, we have devoted considerable effort and space to help you understand and interpret the possible meanings a value may have at one point in time and to help you relate a value at any one time to some baseline reference value for the measure (e.g., your program goals). It is appropriate now, however, to discuss some methods for examining differences in values over time.

Viewed against the yardstick of the length of time most clients remain in treatment, the techniques provided here do not technically reveal "change." Change is a continuous process whose indicator is the existence of significant differences during minute time intervals. You may discover a major difference between the value on a measure for one period and the value on the same measure at a later time. But, if you follow the recommendations in this manual, using calendar months as reporting periods and as the basis for defining cohorts, you still cannot be sure that the value difference between the two times represents a real change in the actual (as opposed to "measured") condition itself. Rather, the value difference may be reflecting the fact that different individuals were being measured at the two different times. Or it may be that conditions at one of the two times were extremely unusual but temporary, not typical of normal program operation (e.g., a successful effort by law enforcement officials to cut off absolutely, if only temporarily, the "supply"), with the result that the value will return to its previous level when the situation outside the program returns to normal.

Since the time intervals between computation of your values will normally be at least as long as 1 month and since the measures suggested in this manual will respond (undesirably, but uncontrollably) to changes in factors outside your program or to client turnover, we refer to patterns of value differences for measures over time as "trends," instead of calling them changes. Trends, in the absence of additional information aside from the measures themselves, can then be considered "tendencies to change (or to remain stable)." Keep in mind that drug treatment is but one event (and perhaps not the most important one) in a client's life.

Directionality of Trends

Of the several ways to indicate value differences over time, we will use two: (1) algebraic differences and (2) graphic presentation. Algebraic differences are simply the arithmetic difference between an earlier and a later value, with plus or minus signs retained to indicate directionality. For example:

Month	Percent unemployed	Algebraic differences
April	48.3	----
May	39.6	-8.7
June	39.6	0.0
July	61.0	+21.4

In this example, the directionality of the trend does not remain constant since unemployment decreases, remains stable, then increases. We determine whether the values increase (+), decrease (-), or remain stable (0.0) by interpreting the algebraic sign of differences, and these terms (increase, decrease, etc.) denote directionality.

Graphically, we may portray the directionality of the trend in the example above by connecting the points plotted for the monthly values with lines, as depicted in figure 7.

In *any* graphical presentation of trends, the measure being considered is plotted along the vertical Y-axis, always with "100 percent" at the top, and *time* is plotted along the horizontal X-axis, with the earliest date always on the left. Then the slope of the line during any time interval (whether lasting 1 month or more) reveals directionality of the trend, a downward slope meaning a value decrease, an upward slope indicating a value increase, and a level slope portraying stability in the values. Of the two methods of indicating value differences over time, graphic presentation probably presents a more obvious indication of directionality.

Although directionality of value differences is the same regardless of measures, *you* must determine for *each* measure the desirability of the respective directions. For example, you may normally want the trend of unemployment in your program to decline, which would be indicated by successive value decreases; but if your community is experiencing a recession, you will probably be doing a "good" job if the trend in program unemployment remains stable, indicated by

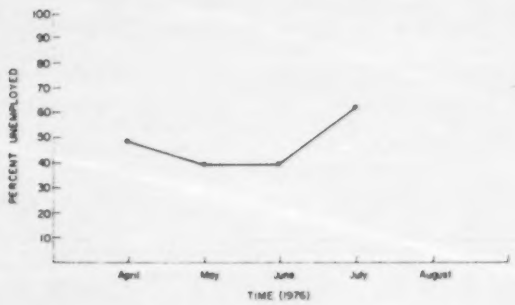

Figure 7. Using graphic presentation to depict directionality of trends.

no algebraic differences between successive values (i.e., a level line on the graph). So the meaning of the directionality of a trend depends on the particular measure being used, the current value(s) for your program (subgroups), and whether you want the value(s) to increase or decrease or remain stable in terms of your program's goals and objectives.

Magnitude of Rate Difference Over Time

Another issue you must decide is whether a trend is moving too quickly or too slowly. In the example in figure 7, the value *increase* between June 1 and July 1 occurred about two and a half times *faster* than did the *decrease* in the value during April. For the unemployment measure, this indicates that the employment picture for the program got "worse" *faster* in June than it had improved during April, in addition to being simply different from the previous month.

The issue of deciding what is acceptable or desirable in a trend is less clear, however, because (1) you do not have *baseline* trends (e.g., for programs nationwide) with which you could compare your program's rates of increase or decrease and (2) *you* must determine for *your* program the tradeoffs in time, effort, and money that are acceptable. For example, with an investment of twice as much money, you may be able to accomplish in 1 month the same-size decrease in the value on a measure (e.g., "Percent Unemployed") as has occurred during the past 4 months; but if you do not have twice as much money, you would still continue to expect the former slow rate of decrease, while a leveling-off of the trend would be a weak indication, and a decline in the successive values for the measure would be a strong indication, of possible problems in your program.

ANALYZING TRENDS

Only one way of analyzing trends is possible using measures whose values are computed on either your entire program or on subgroups. Simply compute the program or subgroup values and determine the difference (algebraically or graphically) each successive month. Figure 8 depicts two trends using the measure "Percent Unemployed" when the values have been computed for male and female subgroups in a program.

Observe that plotting both subgroups on the graph permits comparison of the subgroups'

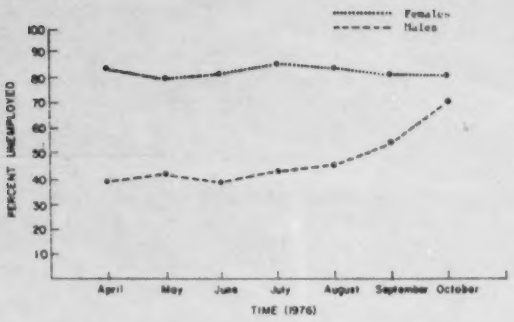

Figure 8. Graphic presentation of two subgroup trends for a measure.

trends with each other, as well as examination of each group's trend separately. The graph in figure 8 indicates that the trend in unemployment for females in the program is fairly stable (whether this is "good" or "bad" depends on whether you have as a program goal reducing unemployment among females). For males, the trend indicates that the unemployment situation is worsening and at an increasingly faster rate (whether this reflects a community economic problem, a lack of program effectiveness, a recent increase in admissions of unemployed male clients, or a combination of all these *cannot* be determined from the graphic information--*you* have to investigate further).

TRENDS FOR COHORT VALUES

Now that you have seen the limited way values on measures computed for your program or for subgroups can be followed over time, this section will discuss two ways to compare trends for values computed for cohorts.

Time-Series Analysis

If we compare each month's admission cohort with previous admission cohorts on a measure, it is possible to determine if the kinds of clients being admitted now are different from those admitted earlier or if new clients have different drug abuse problems. Similarly, we may compare any group of cohorts with identical times-in-treatment to determine whether clients now are being affected differently at the same stage of treatment than were earlier clients. In effect, this comparison of successive cohorts with identical times-in-treatment parallels the comparison of subgroups discussed above; that is, such an analysis subgroups clients on the basis of length of time-in-treatment and thus controls for the effects of time. The unique feature of this kind of analysis is the subsequent step of comparing the resulting "series" of successive identical subgroups.

Figure 9 depicts time-series trends for four admission cohorts. The fact that each series of longer time-in-treatment cohorts generally has lower "Percent Unemployed" values raises at least two possibilities: either (1) the program is becoming increasingly effective in improving the employment situation of clients who stay in treatment longer, though there is little change in its effect on clients who have been in treatment for only two months, *or* (2) clients who cannot find employment soon after admission tend to drop out, leaving a higher proportion of clients who have jobs in treatment. This last possibility, if true, suggests that both the program's effects and the retention of clients might be improved if more emphasis were placed in the early stages of treatment on helping clients get jobs.

Another advantage of performing time-series analysis in this way is to obtain additional information that can be used in interpreting program trends. You can identify whether the trends for the program are results of different types of clients being treated now than had been treated earlier. For example, if you look at the trend in unemployment ratios for successive 1-month cohorts (dotted line), another question can be addressed. The possibility discussed above (with figure 8) that the worsening employment trend for males might be due to an increase in recent admissions of unemployed males can be partially discounted. The time-series analysis in figure 9 reveals

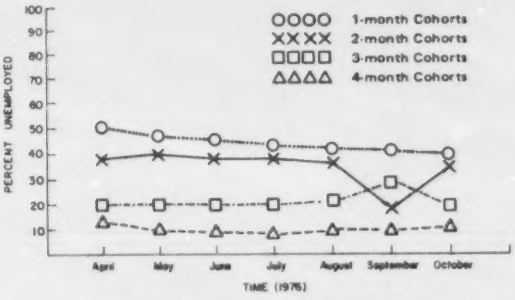

Figure 9. Time-series analysis of treatments for successive cohorts with identical times-in-treatment to compare different time-in-treatment trends.

that the unemployment rate has been relatively stable for all successive cohorts of new clients. A time-series type of analysis only for males within each successive 1-month cohort would be required to *completely* discount the possibility.

"True" Cohort Analysis

The final method of dealing with trends for cohorts is, in many ways, the simplest and most logically appealing technique in this manual. You already know that the best way to assess an individual's progress in a program is to follow his career through the program, reflected in a clinical record. But such detailed individual information, given for many individuals, is not too informative about an overall program. Yet program measures, especially when computed only at the program level, ignore some important individual variations. "True" cohort analysis allows you to use measures which tell you something about your program, while at the same time permitting certain important characteristics of individuals to be taken into account.

The graph in figure 10 depicts the same series of cohorts already portrayed in figure 9. But the lines connecting points represent cohort analysis of unemployment ratios for four admission cohorts, each of which is followed through four months of its "career" in the program.

If you recall the discussion earlier about the constancy of the membership of a cohort, you can see now that each trend plotted in figure 10 reflects value differences over time for the *same* individuals (except for terminations). Hence, we can be confident that for clients admitted during April 1976 (dotted line) the employment situation is improving. We do *not* have to investigate whether this improvement is a result of new clients entering the program who are already employed, since new admissions cannot get into this cohort. It is possible that the improvement is a result of unemployed clients terminating, leaving the cohort with members most of whom are employed; this possibility should be investigated by examining the employment status of clients dropped from this cohort. Whether the *rate* of the decrease in unemployment is satisfactory is a judgment the cohort analysis cannot make--you must make that decision; but the analysis does identify for you that the trend is a decrease rather than an increase in unemployment and how fast that decrease has been occurring.

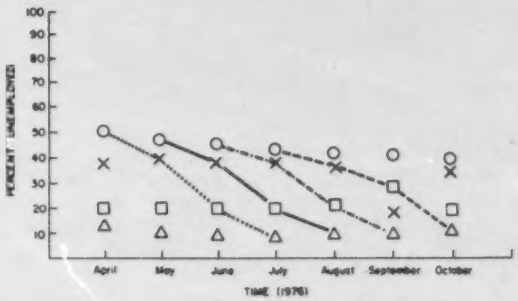

Figure 10. "True" cohort analysis tracing careers of four cohorts through treatment.

CHAPTER VIII
ADDITIONAL PROCEDURES

DEALING WITH SMALL GROUPS

Now that you are familiar with computing measures, subgrouping, constructing cohorts, and the examination of trends, we can illustrate the discussions earlier (see pages 40 and 49) about minimum size of the groups for which you compute measures. Most of our examples thus far have disregarded the problem of minimum size of groups so that the presentation of other points would be relatively uncomplicated.

For example, the dotted line in figure 11 shows the cohort analysis of a fictitious April 1976 admission cohort's "Percent Unemployed" values and is quite regular. But many smaller programs will have only a few clients in each cohort or subgroup and even larger programs will encounter this problem for detailed subgroupings or for cohorts with long times-in-treatment. The dashed line in figure 11 is therefore a more realistic portrayal of the situation when the number of clients in a subgroup or cohort becomes small.

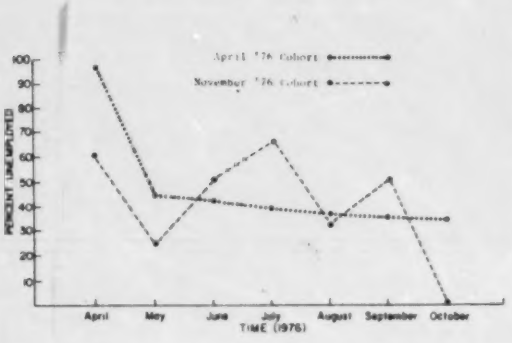

Figure 11. An illustration of the lack of reliability of measures computed on small-sized groups

For the November 1975 cohort, the dashed line results from the following conditions:

Month	Total number of clients in cohort	Number of cohort members unemployed	"Percent unemployed"
April	5	3	60
May	4	1	25
June	4	2	50
July	3	2	67
August	3	1	33
September	2	1	50
October	1	0	0

As you can see, with such small groups the "trend" lines are mostly a consequence of the disproportionately great effect a change in a single client's status has. Treatment efforts may indeed cause the change in client status, but it must be kept in mind that the massive differences in values result at least partially, and perhaps only, because of the arithmetic. Obviously, then, each month's values for such small groups must be suspect; furthermore, analysis of trends of monthly values for these small groups is probably useless since changes over time become increasingly an illustration of the arithmetic differences and less a result of program efforts.

Once you determine that your subgroups or cohorts have become so small that values computed from month to month seem erratic and unrelated, you should attempt to use the methods discussed earlier (see pages 40 and 49) to compensate for the fluctuations resulting from the arithmetic of small numbers. In addition, it is possible to use a variation of the method of "moving averages" to adjust the

monthly values of small subgroups or cohorts so that smoother (and more meaningful) trend lines can be plotted. This technique involves computing several months' values on a measure for each subgroup and/or cohort, no matter how small, in the normal way, as illustrated for figure 11. Then an average is computed for each overlapping set of three consecutive monthly values. Using the monthly "Percent Unemployed" values in the example for figure 11, we add together and average the values for the first three months (i.e., for April, May, and June), and place this average, 45, at the midpoint of that time span, the month May. This is the first of our three-month averages.

The second three-month set of values is created by dropping the previous earliest month's value, then using the remaining two months' values and the next consecutive month's value. Thus the value for April would be dropped and the value for July would be added. Again, these three figures would be averaged, and the resulting three-month average value would be plotted for the month which is midway in the three-month period. Successive sets of three months, each overlapping two months of the preceding set, are averaged in this way until the last month, October in our example, is reached, as shown below:

Month	Actual monthly "percent unemployed" values for the Nov. 1975 cohort	Three-month sets of raw values used for computing moving averages
April	60	60
May	25	25 25
June	50	50 50 50
July	67	135 67 67 67
August	33	142 33 33 33
September	50	150 50 50
October	0	150 0
		83
Moving Average "Percent Unemployed" Values →		45 47 50 50 28

We suggest that you actually plot the values above on the graph in figure 11 to compare the "saw-toothed" pattern of the dashed line (raw values) with the smoother line that results from the three-month moving averages. Just examining the figures above for the example moving averages, however, reveals several important points. Notice first that the two months at the extreme ends of the total period do not have three-month averages, because they are not the middle month for a three-month period. Secondly, using the moving average technique is necessary because the raw values for small groups fluctuate so much that no trend of program effects can be discerned. It can be seen from the example moving average values, even before plotting, that the extreme fluctuations in the raw values have been eliminated. And while no trend was evident from the original raw values, even when they are plotted as the dashed line in figure 11, we can tell immediately from the moving average values that unemployment increased for the November 1975 admission cohort during the first part of the period covered, April through October, then remained constant for a while and improved quite considerably toward the end of the period. Unfortunately, adjusting the raw values with the moving average method does not tell us whether the resulting trend results from program efforts or from nonprogram factors, nor does it tell us the extent to which these different sets of factors are important in combination. As discussed earlier, only by further investigation can the answers to these questions be determined. But by reducing the effects of the complications arising from the arithmetic of small numbers, the moving average technique at least reveals what the trends in your outcome measures have been and makes possible investigation of the causes.

SOME FURTHER SUGGESTIONS

We have suggested that you use your own knowledge of your program, information you may already collect, and some ingenuity to devise custom fit measures that will make your self-evaluation work more rewarding. At least three other steps can be taken to adapt the instructions in this manual to your program and to help make the entire evaluation process more responsive to the needs of your program and, more importantly, of your clients.

Staff Feedback

Earlier (see p. 3) we suggested that you encourage participation of the clinical staff in the data collection process. Now we strongly urge that you publish results of computation of the measures for distribution to your staff so that they also can be involved in all phases of the evaluation process and see the benefits of the additional work required of them for self-evaluation. Of course, staff members can interpret and make use of the evaluation results only if they have become familiar with the meanings and limitations of the measures. Furthermore, it is suggested that you regularly meet with your staff members for the express purpose of getting their opinions about the worth of the individual measures, their suggestions about ways to improve the measures, and information from them which might be helpful in interpreting the measures.

Adapting the Information for Machine Processing

Once information for computing the measures is collected, the time and effort involved in performing the actual calculations can be reduced significantly if some kind of automated machine data processing is available. A huge computer with special programs is not necessarily the only method you can use; regularly recording all the information on a worksheet similar to one of those printed in Appendix A or B and using an adding machine is also a way to standardize data collection and reduce the time and effort required for both data collection and computation of the measures. However, it is suggested that you seek technical assistance from NIDA; from your State's drug abuse authority; or from local sources such as an umbrella drug abuse agency, your city government's data processing department, or data processing experts employed by computer firms, business colleges, and your State's university or community college system. With the relatively widespread availability of standard packages of computational programs and of data processing equipment (such as key-punch machines, counter-sorters, and even computers of various sizes), the instructions and measures in this manual can be easily, quickly, and inexpensively adapted to machine processing.

Client Involvement in Evaluation

Finally, you may find it surprisingly productive to involve clients actively in some phases of the evaluation process. While you need not restrict or alter your evaluation process solely because clients want you to, their suggestions may be useful in revealing deficiencies, needed adjustments, or factors not considered in this manual, all of which might improve your self-evaluation. A possible way to accomplish client involvement is to invite all clients to meet with you and your staff annually or semi-annually for reporting the period's evaluation results and to allow clients to respond critically. At a minimum, you will learn information about the clients and their views on treatment and program operation which may prove useful in interpreting subsequent values on the measures of program effectiveness.

SUGGESTED READINGS

The following are a few references that you may find helpful. They all go into considerably more detail about various issues associated with the evaluation process. Single copies of documents available from the National Clearinghouse for Drug Abuse Information (NCDAI) can be obtained free of charge on request.

Blair, Louis H., and John Sessler
1974 *Drug Program Assessment*, Washington, D.C.: The Drug Abuse Council.

Caro, Francis G. (Ed.)
1971 *Readings in Evaluation Research.* New York: Russell Sage Foundation.

Drug Use in America: Problem in Perspective
1973 Second Report of the National Commission on Marijuana and Drug Abuse. U.S. Government Printing Office, No. 526-00003.

Evaluation Handbook, Second Edition.
1972 Second printing by the Office of Program Evaluation. Washington, D.C.: Agency for International Development, M.C. 1026.1, Supplement II.

Greene, Mark H., and Robert L. Dupont (Eds.)
1974 *The Epidemiology of Drug Abuse.* Washington, D.C.: National Institute on Drug Abuse. U.S. Government Printing Office, No. 041-010-00019. (Available from NCDAI.)

Program Evaluation in the State Mental Health Agency: Activities, Functions, and Management Uses.
1976 Rockville, Maryland: National Institute of Mental Health. U.S. Department of Health, Education, and Welfare, Publication No. (ADM) 76-310.

Quick Evaluation Methodology
1973 Special Action Office for Drug Abuse Prevention. Monograph Series A, Number 2. (Available from NCDAI.)

Sells, S. B. (Ed.)
1974 *The Effectiveness of Drug Abuse Treatment*, Vols. I and II. Cambridge, Massachusetts: Ballinger Publishing.

Suchman, Edward A.
1967 *Evaluative Research: Principles and Practice in Public Service and Social Action Programs.* New York: Russell Sage Foundation.

Struening, Elmer L., and Marcia Guttentag (Eds.)
1975 *Handbook of Evaluation Research*, 2 volumes. Beverly Hills, California: Sage Publications, Inc.

Weiss, Carol H.
1972 *Evaluation Research.* Englewood-Cliffs, New Jersey: Prentice-Hall Publishing Co.

1972 (Ed.) *Evaluating Action Programs: Readings in Social Action and Education.* Boston, Massachusetts: Allyn and Bacon, Inc.

Why Evaluate Drug Education?
1975 Task Force Report of the Prevention Branch of the Division of Resource Development, National Institute on Drug Abuse, Washington, D.C.: U.S. Government Printing Office, No. 017-024-00461-8. (Available from NCDAI.)

The following documents are sources of information about other kinds of measures

that might be useful in your self-evaluation efforts.

Greene, Mark H.
1974 *Estimating the Prevalence of Heroin Use in a Community.* Special Action Office for Drug Abuse Prevention, Washington, D.C.: Monograph Series A, Number 4. (Available from NCDAI.)

Heroin Indicators Trend Report
1976 A publication of the Forecasting Branch, Division of Resource Development, National Institute on Drug Abuse, Washington, D.C.: U.S. Department of Health, Education, and Welfare Publication No. (ADM) 76-315. (Available from NCDAI.)

Hunt, Leon Gibson
1974 *Drug Incidence Analysis.* Special Action Office for Drug Abuse Prevention, Washington, D.C.: Monograph Series A, Number 3. (Available from NCDAI.)

WORKSHEETS

U.S. GOVERNMENT PRINTING OFFICE : 1977 O—230-356

WORKSHEET FOR DARP FORMULAS _____ COHORT #_____ (MONTHS IN TREATMENT AS OF _____)
CURRENT MONTH/YEAR

CLIENT NAME	ADMISSION DATE	DISCHARGE DATE	TREATMENT MODALITY				INTAKE DRUG USE PATTERN				RACE - ETHNICITY				AGE GROUP				SEX		EMPLOYMENT					CURRENT DRUG USE				CRIME		TERMINATIONS				OTHER MEASURES								
			MM	TC	DFOP	DTOP	DA-OP	DA-OP+	LDA-OP+	NON-OP	B	PR	MA	W	UNDER 18	18-21	22-25	OVER 25	M	F	E_f'	H	S	E_p	U	O	D_{no}	D_l	AA	C	I	T'	S_t'	V_t'	RB_t'	TR_t'	A'	R_o'	L	M	G	J	W_t'			

WORKSHEET FOR CODAP FORMULAS _____ COHORT #____(MONTHS IN TREATMENT AS OF _____)
CURRENT MONTH/YEAR

CLIENT NAME	ADMISSION DATE	DISCHARGE DATE	TREATMENT MODALITY				INTAKE DRUG USE PATTERN				RACE – ETHNICITY				AGE GROUP				SEX		EMPLOY-MENT			CURRENT DRUG USE					CRI-ME	TERMINATIONS					OTHER MEASURES								
			MM	TC	DFOP	DTOP	DA-OP	DA-OP+	LDA-OP+	NON-OP	B	PR	MA	W	UNDER 18	18-21	22-25	OVER 25	M	F	E'_f	E'_p	U'	O'	S'_{1m}+Dep	P'_u	AA'	C'	T'	S'_t	V'_t	RB'_t	TR'_t	A'	R'_0	L	M	G	J	W'_t			